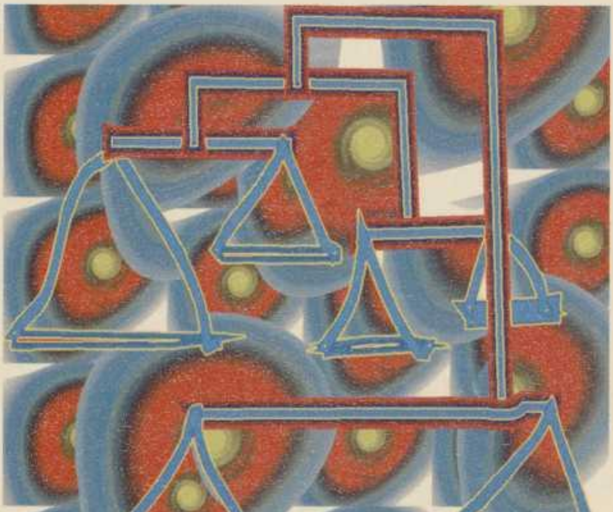
FUNDACION BBV

MÉTODOS CUANTITATIVOS PARA LA TOMA DE DECISIONES (con aplicaciones en el ámbito sanitario)



Daniel Serra de La Figuera

Fundación BBV

La Fundación BBV, alerta siempre a las grandes transformaciones de la sociedad actual, reconoce que uno de los mayores cambios que se han producido en los países occidentales ha sido el experimentado en la atención médica. La medicina tradicional, basada en una relación simple del médico con el enfermo, ha ido dando paso a formas cooperativas y colectivizadas de la asistencia sanitaria, en las que la atención a la salud y los cuidados de la enfermedad constituyen un derecho humano básico y una obligación para toda la sociedad.

El primer Encuentro en el que trabajó la Fundación en materia de Sanidad tuvo lugar en 1992 y sirvió como preparación y antecedente para el diseño de una línea estratégica de investigación sobre Salud y Modelos Sanitarios. Desde entonces, la Fundación intenta contribuir al debate sobre la Sanidad en España desde una perspectiva pluridisciplinar, formulando propuestas de mejora para lograr una mayor eficiencia y equidad del sistema. Esta área estratégica integra tres líneas de investigación:

La Formación de los Profesionales de la Salud analiza la diversidad de las profesiones sanitarias y el desequilibrio existente entre el rápido desarrollo del sistema asistencial y los esquemas y métodos formativos de los profesionales sanitarios, especialmente del médico. Las investigaciones y debates se centran en la situación de las distintas titulaciones, las carencias de las especialidades médicas y de la enfermería y la formación de gestores y otros profesionales sanitarios.

Organización y Gestión del Sistema de Salud estudia las alternativas de organización sanitaria con el objetivo de ofrecer planteamientos de futuro, diseñados a la luz de modernos criterios de eficiencia y equidad. El nuevo papel de los ciudadanos como consumidores de servicios sanitarios, las relaciones entre financiación e incentivos y la eficacia de la gestión son algunas de las cuestiones incluidas en esta línea de investigación.

Salud, Comunicación y Sociedad examina las actividades y conductas de los ciudadanos en relación con la salud, así como la influencia de los programas de comunicación y los medios empleados en la información sanitaria dirigida a la sociedad.

Métodos cuantitativos para la toma de decisiones (con aplicaciones en el ámbito sanitario) presenta algunos de los métodos actualmente más relevantes que ayudan a los gestores a tomar decisiones, tales como la programación lineal y la gestión de colas. Este libro se ha escrito básicamente para dos tipos de usuarios. Por un lado, para el administrador general o administrador en potencia; por otro, para el estudiante de cursos de posgrado en gestión y administración.

FUNDACIÓN BBV
Documenta

FUNDACION BBV

MÉTODOS CUANTITATIVOS PARA LA TOMA DE DECISIONES

(con aplicaciones en el ámbito sanitario)

Daniel Serra de La Figuera

Fundación BBV

Métodos cuantitativos para la toma de decisiones (con aplicaciones en el ámbito sanitario)

Edita Fundación BBV. Documenta
Plaza de San Nicolás, 4
48005 Bilbao
Depósito legal: M. 42.333-1999
I.S.B.N.: 84-95163-25-X

Imprime Sociedad Anónima de Fotocomposición
Talisio, 9 - 28027 Madrid

Métodos cuantitativos
para la toma de decisiones
(con aplicaciones en el ámbito sanitario)

ÍNDICE

PRESENTACIÓN

Este libro se refiere al uso de modelos cuantitativos en la resolución de problemas de gestión y administración de sistemas complejos, con especial énfasis en la toma de decisiones. Ha sido escrito para dos tipos de usuarios: el administrador general o administrador en potencia, que pueden sacar provecho de su uso, o de la comprensión de los modelos cuantitativos, o para estudiantes de cursos de posgrado o masters en gestión y administración. Otro aspecto a destacar es que todos los problemas planteados en el libro están relacionados con la gestión de sistemas sanitarios. La razón es bien simple: la toma de decisiones en el entorno sanitario suele ser muy compleja. Si bien muchos problemas con los cuales se enfrentan los gestores y responsables en organizaciones sanitarias no son analíticamente diferentes de problemas en otras industrias, muchos otros son bastante peculiares debido a algunas de las características intrínsecas de los sistemas de atención sanitaria.

El libro no pretende ser exhaustivo en cuanto a las técnicas *existentes*, ya que existe un sinfín de excelentes manuales de técnicas cuantitativas y de investigación operativa (algunos de ellos se citan al final del libro). La gran diferencia entre este libro y los manuales clásicos radica en que el nivel de complejidad matemática se mantiene al mínimo nivel posible, y se hace especial énfasis en el planteamiento de modelos y en explicar cómo algunas de las técnicas *existentes* pueden ayudar a solucionar problemas que aparecen en cualquier organización. Por ello, para su lectura no se necesita una formación matemática previa; incluso se puede decir que en todo el libro no se utilizan más que las cuatro operaciones aritméticas básicas: sumar, restar, multiplicar y dividir.

En la referencia bibliográfica a este libro, en el website de la Fundación BBV (http.//www.fbbv.es), puede cargarse una aplicación práctica con ejemplos, diseñada por el autor.

Fundación BBV

1. INTRODUCCIÓN: MÉTODOS CUANTITATIVOS EN LA GESTIÓN SANITARIA [1]

El desarrollo de la investigación operativa, según muchos autores, ha representado uno de los avances científicos más importantes desde mediados del siglo XX. Actualmente es una herramienta utilizada en muchos campos de la administración, de la economía y de la ingeniería. Existen muchos libros de texto sobre el tema y miles de artículos científicos en revistas especializadas.

La investigación operativa tiene como base el método científico para investigar y ayudar a tomar decisiones sobre los problemas complejos de las organizaciones de hoy en día. Básicamente la investigación operativa sigue los pasos siguientes: (1) la observación de un problema, (2) la construcción de un modelo matemático que contenga los elementos esenciales del problema, (3) la obtención, en general con la ayuda de un ordenador, de las mejores soluciones posibles con la ayuda de algoritmos exactos y heurísticos y finalmente (5), la calibración y la interpretación de la solución y su comparación con otros métodos de toma de decisiones.

Un ejemplo simple, el problema de la asignación, nos puede servir para ilustrar la dificultad esencial de la investigación operativa. Un hospital tiene 70 trabajadores con cualificaciones diferentes (médicos, enfermeros, ATS, personal de administración, etc.) que hemos de asignar a 70 actividades también diferentes. Si pudiéramos determinar un valor que reflejase la asignación de un trabajador a una tarea determinada, tendríamos que escoger una entre 70! formas posibles de permutación de las asignaciones que maximice el valor total. Como que 70! es aproximadamente igual a 10^{100}, necesitaríamos un ordenador

---

[1] Una versión de este capítulo ha sido publicada en el volumen dedicado a la Economía de la Salud de la revista *Papeles de Economía*, en 1998. La profesora Helena Ramalhinho fue co-autora del artículo.

que ejecutase 1.000.000 de operaciones por segundo durante aproximadamente 10^{87} años (muchas veces la vida proyectada del universo) para examinar todas las permutaciones. Problemas de decisión como éste son muy comunes y se tienen que desarrollar modelos de programación matemática, métodos matemáticos para obtener soluciones a los modelos y algoritmos de ordenador (procedimientos paso a paso) muy eficientes. Se dice que la investigación operativa constituye el 25 % del tiempo total utilizado por los ordenadores para resolver problemas científicos.

La investigación operativa ha tenido un impacto impresionante en la mejora de la eficiencia de numerosas organizaciones en todo el mundo. Existen innúmeras aplicaciones con éxito en todos los campos en donde la toma de decisiones es compleja y que pueden implicar para la organización grandes inversiones o cambios en la organización que determinen su futuro.

La investigación operativa tiene una presencia importante en la gestión y planificación de sistemas sanitarios. Si bien muchos problemas con los cuales se enfrentan los gestores y responsables en organizaciones sanitarias no son analíticamente diferentes de problemas en otras industrias, muchos otros son bastante peculiares debido a algunas de las características intrínsecas de los sistemas de atención sanitaria. Algunas de estas características son la posibilidad de muerte o calidad baja de vida, la dificultad de medir adecuadamente el valor y la calidad de los resultados obtenidos, la toma conjunta de decisiones por diversos colectivos de índole muy diferente (personal médico, de enfermería, gestores), los mecanismos de pago por parte de terceros de los diagnósticos y los tratamientos, y el concepto del acceso a la atención sanitaria como un derecho universal. Este capítulo no pretende examinar exhaustivamente la literatura sobre investigación operativa en el campo sanitario, sino dar una idea de los diferentes ámbitos sanitarios en los cuales se ha utilizado con éxito la investigación operativa para mejorar las decisiones de gestión y mostrar algunas aplicaciones que se han producido en ellos.

1.1. Planificación y diseño de sistemas sanitarios

1.1.1. *Planificación y estrategia*

En la mayoría de los países occidentales, la planificación sanitaria se realiza a nivel nacional o regional. En estos países (quizás con excepción de los Estados Unidos), la interacción entre historia, cultura, economía, política y otros factores han conducido a elaborar estrategias muy similares para proveer acceso universal y control de costes, incluyendo alguna forma de seguridad social, un sistema de pago único para todos los implicados, y techos presupuestarios formales o informales. Debido a estas elecciones estratégicas, el proceso de planificación a nivel nacional o regional determina ampliamente el tipo y los niveles de los servicios y tecnologías y frecuentemente (quizás implícitamente) el nivel del personal sanitario y los niveles de servicios recibidos por la población. La investigación operativa ha colaborado ampliamente en el diseño de estos sistemas. Por ejemplo, en el Reino Unido se desarrolló un sistema de ayuda a la decisión para planificar los recursos necesarios en servicios de tratamiento de SIDA. El modelo incorpora previsión de la demanda por categoría de pacientes, protocolos de atención y necesidades presupuestarias y de recursos para los planificadores centrales y locales pertenecientes al Servicio de la Salud Nacional Británico [2]. También éste se puede utilizar con los modelos de programación matemática de planificación de atención equilibrada [3] incorporados en un ordenador personal para determinar la necesidad de recursos y las asignaciones correspondientes para regiones y distritos sanitarios.

Otro tipo de aplicaciones se encuentra en el diseño de distritos sanitarios en función de la previsión de la demanda por categoría de pacientes y la oferta existente de los servicios sanitarios. El modelo matemático en general tiene como objetivo minimizar la distancia media de acceso a los servicios por parte de la población. En algunos casos en que se están reorganizando los distritos algunos autores han utilizado un segundo criterio: minimizar la desviación entre los distritos existentes y los propuestos, que mejoran la aceptación política. También se han aplicado modelos de programación lineal para regionalizar servicios específicos, como laboratorios, en donde se consideran los costes anuales de operación, el coste por acto, los niveles de utilización y la distancia viajaba por los pacientes. En general, el objetivo consiste en la minimización de los costes totales de operación.

Otro ámbito muy importante dentro de la planificación sanitaria es el de la localización de los diferentes tipos de servicios sanitarios. La buena ubicación de ambulancias, de centros de urgencias, de centros de asistencia primaria, entre otros, es uno de los determinantes principales de la calidad del servicio. Se han desarrollado muchos modelos para poder obtener una serie de localizaciones óptimas en función de los objetivos propuestos. En general, los modelos pueden clasificarse en tres apartados, siguiendo la tipolo-

---

[2] British National Health Service.
[3] Balance of Care (BOC) models.

gía de sus objetivos. El primer grupo corresponde a los modelos de cobertura. Básicamente, se trata de encontrar la localización óptima de un número fijo de servicios de forma a maximizar su cobertura dentro de un tiempo o una distancia estándar predeterminada. El segundo grupo corresponde aquellos modelos que utilizan como objetivo la minimización de la distancia o tiempo de desplazamiento medio de los usuarios potenciales del servicio. Por último, el tercer tipo de modelos tiene como objetivo la minimización de los costes totales del servicio (costes de apertura, costes de operación y costes de desplazamiento). En general, estos modelos se basan en la programación lineal entera. En algunos casos se combinan dos o más objetivos para poder examinar, por ejemplo, el intercambio entre el número determinado de servicios y el grado de cobertura. También se han desarrollado modelos que ubican servicios que tienen una naturaleza jerárquica, tales como centros básicos de asistencia primaria y centros con especialidades ambulatorias. Finalmente, algunos modelos combinan la ubicación óptima de servicios con el diseño de los distritos sanitarios correspondientes.

Muchos estudios se realizan a nivel institucional para planificar la expansión o reducción de servicios y centros. En el campo sanitario, este problema se complica debido a la interdependencia entre servicios y entre instituciones creada por las características propias de la profesión médica, co-morbididades de pacientes, niveles desconocidos de demanda latente y la creación potencial de demanda. Para evitar estos problemas, la mayoría de los estudios, o se concentran en especialidades determinadas (obstetricia, pediatría y psiquiatría) o en grandes modelos generales de planificación. Un ejemplo del primer tipo de estudios consiste en estudiar cómo el hecho de cerrar un determinado tipo de servicio (por ejemplo, obstetricia) en algún centro puede afectar a la carga del sistema y a los resultados de una red hospitalaria. Para ello se suelen modelar tres factores: relación con otros servicios, redistribución de la carga de pacientes y medidas financieras. La relación con los otros servicios determina si el hecho de cerrar el servicio en algunos lugares cambia el nivel general del servicio medido en términos de mano de obra, necesidades de camas, número de salas de parto, y medidas de resultados (que incluyen en nuestro ejemplo mortalidad neonatal, número de nacimientos, porcentaje de cesáreas). El modelo también estudia cómo el volumen de los pacientes se redistribuye entre las unidades restantes después de cerrar algunas y cómo esto afecta al nivel general de servicio en estos centros. El análisis financiero determina cómo los costes directos, específicos a un determinado servicio y marginales, pueden ser utilizados para obtener el ahorro neto a la hora de cerrar algún servicio y los costes adicionales al aumentar la carga en otros centros.

También es muy frecuente la aplicación de la investigación operativa para determinar la apertura de nuevos servicios. Por ejemplo, examinar el efecto de abrir una nueva unidad determinada en un hospital. La pregunta que en general se intenta responder es cómo redistribuir la capacidad existente del hospital para poder crear esta nueva unidad, en términos de camas, personal, etc. Para ello es necesario conocer el flujo de llegadas al nuevo servicio y la duración del tratamiento. En estos casos se utilizan modelos de gestión de colas y simulación con el objetivo de encontrar el número óptimo de camas necesario para esta nueva unidad.

La simulación también se utiliza para examinar si un nuevo servicio puede mejorar la productividad, la calidad o los resultados económicos de un centro o de un sistema sanitario. Por ejemplo, varios modelos de simulación se han aplicado para estudiar los efectos de modificaciones en el tipo de atención sobre la productividad del sistema (e.g., sustitución de médicos por asistentes o médicos por enfermeras en atención primaria). Estos modelos de simulación utilizan la programación lineal como herramienta básica y se utilizan parámetros tales como las horas de funcionamiento, los recursos disponibles, los horarios de visita de los pacientes y su proceso de llegada, los tiempos y las necesidades de servicio y la secuencia de las tareas para adecuarse a las necesidades de servicio de los pacientes. En general, este tipo de estudios se ha mostrado más eficiente que el tradicional uso de las funciones de producción para determinar la productividad, ya que incorporan en el análisis muchos más parámetros sobre el funcionamiento del sistema. Un ejemplo de ello fue la aplicación de la simulación para estudiar los efectos sobre la productividad de la incorporación de residentes a la atención médica. Se demostró que, con la incorporación de los médicos residentes, se atenderían un total de 393 pacientes por semana, comparado con 208 sin ellos, y que la productividad neta aumentaría a lo largo del tiempo, resultado que también sugería un estudio realizado con funciones de producción. Sin embargo, el modelo de simulación iba más lejos, ya que debido a que el tiempo de supervisión médica y la congestión aumentarían durante los períodos iniciales con los residentes, habría un deterioro significante en el tiempo de espera de los pacientes sin importantes aumentos en los resultados. Los autores concluyeron que, aunque a largo plazo el uso de médicos residentes aumentaría la productividad del sistema, los costes a corto plazo podrían hacer prohibitiva su utilización, debido a las pérdidas en el número de pacientes y a los costes de supervisión.

1.1.2. *Previsión de la demanda*

La previsión de la demanda agregada y la previsión de censos diarios son un factor importante para mejorar la eficiencia en la utilización de los recursos disponibles para la atención sanitaria. La investigación en sanidad suele utilizar algoritmos bien conocidos en otros sectores económicos para realizar previsiones, más que realizar nuevas aportaciones metodológicas. Esto es debido a que muchos modelos de planificación como los descritos en el apartado anterior necesitan como input la previsión de la demanda. En efecto, la previsión de la demanda es un componente importante de la utilización de la capacidad y de la calidad del servicio en hospitales, centros de atención primaria, centros de urgencias y casi podría decirse que en cualquier componente de lo que es la asistencia sanitaria. La previsión de la demanda juega dos papeles: por un lado, la determinación de necesidades agregadas para los servicios (el lado de la demanda) y, por el otro, en la planificación censal (el lado de la oferta). El análisis de la demanda diaria determina las decisiones hospitalarias de amplio espectro, que incluyen decisiones sobre asignación de personal, servicios auxiliares, asignación de horarios de pacientes y servicios de apoyo (limpieza, comida, ropa de cama, etc.), y también decisiones que afectan a servicios sanitarios tanto de nueva creación como existentes. En este sentido, la previsión de la demanda en la atención sanitaria se parece mucho a la estructura de problema de planificación agregada en entornos industriales y manufactureros.

Existen varias técnicas de previsión de la demanda. En primer lugar, hay las aproximaciones cualitativas, tales como el análisis histórico, que utiliza el estudio de entornos similares o la historia de la propia organización para determinar la demanda futura. Esta técnica, barata y fácil de utilizar, ignora cambios ambientales y no permite analizar nuevas decisiones en un entorno diferente al conocido. Otra aproximación cualitativa es la técnica Delphi, con la cual se obtienen predicciones por un grupo de expertos a través de diversas reuniones y debates. El proceso de discusión se repite hasta llegar a un consenso. Esta técnica está sujeta a sesgos ideológicos que dificultan su utilización en entornos en donde las predicciones pueden no ajustarse a la visión estándar de las operaciones del sistema. Las otras áreas de previsión de la demanda utilizan una aproximación cuantitativa. En general, se basan en técnicas de regresión tales como medias móviles, mínimos cuadrados ordinarios y modelos autorregresivos tipo ARIMA [4] de series temporales.

---

[4] «Auto-Regressive Integrated Moving Average».

La diferencia básica entre estas técnicas, tanto cualitativas como cuantitativas, reside en el grado en el cual el juicio subjetivo influencia el modelo. Por ejemplo, la técnica Delphi intenta formalizar juicios explícitos de expertos. Otras técnicas, como los modelos econométricos, contienen juicios implícitos (como la especificación del propio modelo). El impacto de estos juicios, sean implícitos o explícitos, sobre los resultados de la previsión puede ser importante. Por ello, al escoger el modelo de previsión, se tendrían que considerar estos juicios, y los resultados se deberían comparar con los de otras aproximaciones. Del mismo modo, sería necesario el realizar análisis de sensibilidad para valorar la robustez del modelo.

1.1.3. *Planificación de la capacidad*

La planificación de la capacidad está relacionada con las decisiones que afectan los niveles apropiados de centros, equipos y personal para atender una determinada demanda. En la atención sanitaria, la planificación de la capacidad en general se centra en decisiones tales como el número de camas, la capacidad de los quirófanos, la asignación de la capacidad a diferentes servicios, del equipamiento en general, de servicios auxiliares, así como de factores que afectan a la utilización de la capacidad, tales como los flujos de pacientes, los niveles de personal médico y auxiliar y las combinaciones de personal con diferentes calificaciones. Al examinar estos factores, los proveedores buscan cómo aumentar la productividad de los recursos existentes y la mejora de la calidad del servicio. En general, el objetivo es la maximización de la utilización de recursos de equipos de capital o la minimización de las ineficiencias y atrasos en la prestación de servicios sanitarios.

Muchos estudios permiten el proceso de planificación de la capacidad en un sistema de colas y utilizan la simulación para obtener resultados. Por ejemplo, para determinar la asignación de camas dentro de un hospital en general se suele realizar un modelo de simulación que incorpora las categorías de los pacientes en función de su diagnóstico, sexo, tipo de ingreso, tiempo de la demanda, tipo de habitaciones y tipo de médico (principal o residente). Los resultados del modelo ofrecen datos sobre porcentaje de ocupación, censo medio diario y número anual de pacientes-año. Otros datos secundarios indican las peticiones de ingreso, reservas, rechazos debido a no encontrar un día de ingreso adecuado, listas de espera y transferencias. El simulador también puede permitir la evaluación de diferentes reglas de asignación de pacientes.

Otras técnicas utilizadas en el diseño de la capacidad son los procesos de Markov y Semi-Markov, sobre todo en sistemas en donde el tipo de atención del paciente cambia en función de su estado de salud. En este tipo de técnicas se calculan las probabilidades transicionales de pasar de un estado a otro para obtener el nivel esperado de utilización y la consecuente capacidad del sistema. Si se conoce la tasa de llegada en equilibrio, el tiempo medio dentro de cada unidad de atención, y la probabilidad de que una unidad determinada esté a plena capacidad, entonces se pueden obtener las tasas de utilización y de servicio del sistema.

1.1.4. *Evaluación de la eficiencia*

La investigación operativa se utiliza con frecuencia para evaluar la eficiencia productiva de un determinado sistema. En general, se define una función frontera de un conjunto de producción formado por todas las combinaciones posibles de factores de producción y de producto final. Se intenta determinar para cada organización dentro del sistema su posición en términos de eficiencia respecto a la frontera o, más específicamente, su distancia respecto a ella. Las distancias más comunes son o en términos de los factores de producción, es decir, la cantidad de factores utilizada con respecto a la que sería suficiente, dado un nivel de producto, si la actividad fuese eficiente; o en términos de producto, que es la cantidad máxima de producto que se alcanzaría si la unidad fuera eficiente, dado el nivel de factor consumido. Cuando la frontera se determina sin una función de producción específica, pero con unas propiedades formales que satisfacen los puntos del conjunto de producción (como por ejemplo la libre disponibilidad de inputs y outputs), se utiliza una aproximación no-paramétrica, que en general utiliza la programación lineal como herramienta para determinar la frontera. Por otro lado, a veces se especifica una forma funcional con parámetros constantes, estimándose sus parámetros de manera que las observaciones queden sobre o por debajo de la función. Con respecto a esta función se mide la eficiencia, que será distinta según la forma funcional especificada *a priori*. Esta es una aproximación paramétrica. El avance de la investigación en este campo en los últimos años ha sido enorme y existe un sinnúmero de aplicaciones para determinar la eficiencia de un determinado servicio en concreto o de los centros hospitalarios en general pertenecientes a una red sanitaria.

1.2. Gestión de operaciones

1.2.1. *Sistemas de información sobre la gestión*

Hoy en día, el nivel de complejidad de las operaciones dentro de cualquier organización tiende a ser muy elevado. La gestión sanitaria no se escapa de esta complejidad, sino que la aumenta. El volumen de información necesario es muy grande, y abarca tanto la necesidad de controlar la gestión económica como la necesidad de generar información médica que puede ser utilizada para mejorar la calidad del servicio. Actualmente se han desarrollado muchos modelos de información sobre la gestión en el campo sanitario. En general, estos sistemas se combinan con algunas herramientas de ayuda a la decisión para mejorar el grado de reacción de las decisiones, la calidad de la atención y la productividad. Las mejoras en la atención pueden obtenerse debido a la reducción en el tiempo de espera de la ejecución de las acciones a realizar determinadas por el médico y la eficaz disponibilidad de los resultados obtenidos, la eliminación de servicios innecesarios, la reducción de errores, y un aumento en la satisfacción de los pacientes. Se consiguen aumentos en la productividad gracias a la disponibilidad de datos para la decisión sobre asignación de personal, horarios, utilización de equipos y por la eliminación de acciones y sistemas redundantes.

1.2.2. *Gestión de pacientes*

El control de la demanda de servicios a través de la correcta programación de los pacientes puede ser un método muy eficaz para ajustar la demanda a los servicios disponibles. Se puede decir que la programación de pacientes en clínicas para servicios externos e internos es una de las aplicaciones más antiguas de la investigación operativa para mejorar la atención sanitaria. Desde los años 50, la investigación operativa ha estudiado extensamente este problema.

En el caso de pacientes externos, la metodología más utilizada es la de gestión de colas junto con simulación. La mayoría de los modelos utilizan como parámetros principales el tiempo medio de consulta, el coeficiente de variación de ésta, la desviación estándar de la puntualidad de los pacientes y el número total de citas para obtener resultados sobre el tiempo medio de espera, el tiempo durante el cual un servicio está desocupado y las horas extras del personal médico. El objetivo consiste en encontrar una organización del sistema (número de médicos y personal auxiliar, organización de la cola, tiempo de asistencia) que minimice el tiempo de espera de los pacientes mientras se mantienen los servidores ocupados. Actualmente se utilizan tres tipos de modelos de programación de pacientes externos: programación por bloques, programación modificada por bloques y programación individual. En la programación por bloques, un número determinado de pacientes es citado dentro de un mismo intervalo horario (por ejemplo, de 9:00 a 13:00) y son atendidos con sistema FIFO [5]. Este sistema suele generar un largo tiempo de espera, alta congestión y tiempo libre mínimo para el personal. La programación modificada por bloques reduce los intervalos horarios. Tiene algunas características muy parecidas al sistema de bloques, pero reduce el tiempo de espera de los pacientes. Por otra parte, el sistema individualizado de asignación de horarios, en general, consigue menores tiempos de espera que los otros dos sistemas, sobre todo si las citas previstas no son significativamente menores que el tiempo de servicio. Sin embargo, para este sistema se necesita mucha información sobre las necesidades y las enfermedades de los pacientes, y a veces es necesario el realizar clasificaciones por parte del encargado de la agenda, y sucesos imprevistos pueden causar serios problemas en ésta.

En el caso de la gestión de citas para pacientes internos, los modelos suelen ser más complejos. El problema tiene tres dimensiones: primero, el programar diariamente los ingresos selectivos junto con los provenientes de urgencias dentro de cada unidad del hospital; en segundo lugar, el agendar los pacientes internos a las unidades apropiadas de tratamiento y diagnóstico durante su estancia en el hospital, y en tercer lugar la programación de las altas hospitalarias y de las transferencias a otros centros. Es obvio que estas tres actividades están ligadas entre sí y dependen mucho de las características de los pacientes y del hospital. En general, la metodología utilizada se basa en las cadenas de Markov, programación matemática, sistemas expertos, heurísticas y simulación. Existen otros métodos más tradicionales asociados con la rutina de la organización. Debido a la información precaria sobre la previsión de la demanda, sobre el funcionamiento interno y a la complejidad antes descrita, la mayoría de hospitales utilizan métodos informales o *ad hoc*, que conllevan a una mala planificación del tamaño de los servicios, la asignación de personal y la utilización de la capacidad, provocando serias ineficiencias en forma de horas punta y horas valle de utilización. Las horas valle suelen crear problemas de exceso de contratación. Las horas punta provocan congestión y, debido a las características intrínsecas del personal de un hospital, no es fácil corregir esta congestión con contratación temporal. Todos estos factores conllevan a una reducción de la calidad del servicio y a un aumento de costes. Estos problemas pueden ocurrir en todos los niveles de la institución. Por otro lado, la utilización de métodos

---

[5] «First In, First Out», el primero que llega es el primero en ser atendido.

cuantitativos analíticos puede mejorar muchos de estos problemas y conseguir una utilización más efectiva y eficiente de los recursos disponibles.

1.2.3. *Gestión y planificación de recursos humanos*

La gestión de los recursos humanos juega un papel primordial en el día a día de los sistemas de atención sanitaria. Esto es debido, entre otros factores, a que el coste de personal suele representar la partida más grande dentro de los costes de operación del sistema. De ahí que, cada vez más, los gestores se concentren más en buscar una mejor administración al enfrentarse a presupuestos más restrictivos. Así como en otras organizaciones, la habilidad de asociar los recursos humanos a la fluctuación de la demanda afecta directamente la eficiencia en las operaciones y a la calidad del servicio. El desarrollo de nuevos métodos para gestionar el personal médico y de enfermería puede contribuir en gran parte a la reducción de los costes en la atención sanitaria.

Si bien, en principio, la gestión de recursos humanos en el ámbito sanitario puede parecer un problema similar al de cualquier organización (uno necesita las personas adecuadas en el sitio adecuado y en su momento adecuado), esto no siempre es así, debido a varios factores que complican seriamente el problema. En primer lugar, existe una interrelación entre los diversos colectivos altamente especializados y cualificados que tienen que estar disponibles en un momento determinado para pacientes diferentes. Estos colectivos incluyen personal de enfermería, de rehabilitación, médicos, técnicos y otros. El personal tiene que estar disponible veinticuatro horas al día, considerando las fluctuaciones de la demanda. También este tipo de personal tiene sus preferencias horarias y condiciones de trabajo. Muchos de ellos también funcionan como profesionales independientes, lo que dificulta su flexibilidad horaria. En segundo lugar, en general es muy difícil medir la calidad del trabajo, lo que provoca problemas cuando se pretende encontrar la combinación correcta de los profesionales necesarios para una atención adecuada. En tercer lugar, en la mayoría de las carreras profesionales existe una estructura organizativa plana con pocas posibilidades de mejora. Como consecuencia de ello, la gestión de recursos humanos tiene que estar enfrentándose continuamente a mantener la satisfacción del personal durante años e incluso décadas, para evitar costes de absentismo laboral, rotación de personal y reducir el descontento.

La mayoría de los estudios basados en la investigación operativa se han concentrado en la gestión del personal de enfermería. Otras aplicaciones se encuentran en la gestión del personal de laboratorio, el departamento de urgencias, de terapia respiratoria, y en el ámbito de los sistemas de sanidad gestionada (HMO [6]). Una vez más, las técnicas más utilizadas incluyen la programación matemática, la gestión de colas y la simulación. Por ejemplo, una manera de gestionar la enfermería es organizarla en una jerarquía de tres niveles de decisión que operar en períodos de tiempo diferentes y con precisión también diferente. Estos son: asignaciones correctivas, diseño de turnos y planificación del personal. Las asignaciones correctivas se realizan diariamente, el diseño de turnos de los enfermeros se diseña para un período de entre cuatro y ocho semanas, la planificación de personal se suele hacer a un año vista.

En un determinado día y dentro de un turno, la capacidad de personal entre las unidades puede ajustarse a fluctuaciones de la demanda y a absentismo no previsto. Estas asignaciones correctivas dependen de las preferencias individuales de los empleados, de su disponibilidad y de su capacitación. El diseño de turnos consiste en asociar la disponibilidad de personal a la carga de trabajo esperada en las diferentes unidades. Para cada empleado se determina los días de trabajo, de descanso y la rotación en un período determinado. En este caso, las preferencias individuales y las necesidades tienen que ser consideradas para evitar la insatisfacción. Estos dos niveles de decisión son tácticos, ya que conciernen a la utilización de un personal ya contratado por la organización con sus conocimientos y especialidades. El tercer nivel corresponde al de las «grandes decisiones», en el que se determinan las necesidades a medio y largo plazo de personal por categorías y especialidades. La mayoría de los modelos estudian los dos primeros niveles. En muchos casos se utilizan modelos econométricos de previsión para predecir el diseño de turnos y las necesidades de horas de personal por unidad. Otros modelos se basan en la gestión de colas y en la combinación de ésta con programación lineal entera mixta, esto es, modelos lineales de optimización, en los cuales algunas de las variables sólo pueden tomar valores enteros.

1.3. Gestión médica

1.3.1. *Cribaje de enfermedades*

La introducción y mejora de tests para detectar enfermedades ha implicado un serio avance en el diagnóstico médico y en la detec-

[6] «Health Maintenance Organizations», Organizaciones para el Mantenimiento de la Salud.

ción de patologías. Estos tests se pueden aplicar individual o colectivamente a un sector de la población. Estos dos tipos de cribaje requieren de una modelización diferente porque implican objetivos diferentes para los decisores, y las restricciones, parámetros y variables del modelo pueden cambiar considerablemente. Por ejemplo, en los cribajes poblacionales, el objetivo puede ser la minimización de la prevalencia de una enfermedad contagiosa en una población sujeta a las restricciones de recursos. Para un individuo, el objetivo puede ser la prolongación de la vida (o de su calidad) y las restricciones pueden estar asociadas a su capacidad o disponibilidad de pago y los parámetros de las características individuales de los pacientes. En cada caso, los decisores son también diferentes: a nivel colectivo puede ser el director de un servicio de salud o gestor de salud pública de una región, y a nivel individual el propio paciente o su médico.

Los programas de cribaje masivos se han utilizado en muchas áreas médicas, tales como el SIDA, hepatitis A, B, no-A, no-B, tuberculosis, sífilis y otras enfermedades infecciosas. En algunos países existen protocolos de cribaje tanto para enfermedades contagiosas como no contagiosas para realizar programas nacionales de salud. Por ejemplo, en el Reino Unido existe un protocolo de cribaje poblacional para la detección de cáncer de cuello de útero, destinado a reducir la prevalencia de la enfermedad y a salvar vidas.

Uno de los problemas de estos cribajes masivos es su coste total. Por un lado, tenemos unos costes directos asociados, tales como el personal y el material necesarios para administrar el cribaje. Por otro lado tenemos unos costes indirectos que pueden ser muy elevados, tales como la inconveniencia y la incomodidad posible del propio test, el coste de falsos positivos que implica la realización de cribajes secundarios y un estrés emocional, e incluso el riesgo de daño físico sobre la persona (por ejemplo, el efecto acumulado de exposición de rayos X o daño por una cirugía innecesaria). Por lo tanto, al diseñar un cribaje hay que tener en cuenta, entre otras cosas: (1) el intercambio entre el coste del test, que aumenta tanto con la frecuencia de su aplicación como con la precisión del test utilizado; (2) el instrumento de aplicación y las tecnologías existentes; (3) la frecuencia de aplicación; (4) el tamaño de la muestra; (5) los problemas logísticos asociados a la implementación del cribaje y (6) la prevalencia de la condición a cribar.

Los estudios epidemiológicos son una de las áreas de investigación más desarrolladas dentro del diseño de cribajes. Estos estudios se preocupan de desarrollar modelos descriptivos de los procesos de la enfermedad, su progresión y los factores causales. En el caso de enfermedades contagiosas, se han desarrollado modelos estadísticos y matemáticos (de ecuaciones diferenciales) para describir el crecimiento, la madurez y el declive de varias enfermedades específicas. Otro campo de investigación asociado al cribaje es el estudio del mantenimiento, es decir, la modelización de la toma de decisiones sobre la inspección y conservación de unidades o sistemas sujetos a deterioro u obsolescencia. En este caso se aplican modelos que incluyen la programación matemática y la simulación.

Los modelos de cribaje individual incorporan, frecuentemente, información sobre la progresión de la enfermedad e intentan minimizar la diferencia entre el tiempo esperado de detección de la enfermedad y su aparición, o maximizar la diferencia entre el tiempo de detección a través del cribaje y el tiempo de auto-detección o de aparición de síntomas evidentes de la enfermedad. Estos modelos tienen una estructura temporal para el individuo, porque éste, a lo largo de su vida, está sujeto a probabilidades condicionales diferentes de que aparezca la enfermedad. Por lo tanto, la programación temporal de los cribajes se determina secuencialmente con la utilización de resultados anteriores, o simultáneamente para toda la vida, por ejemplo, cada x años. Muchos estudios han sido elaborados para obtener un uso eficiente de los cribajes a nivel individual. Se ha demostrado que, en general, utilizando la etiología y el progreso de una enfermedad y su relación con la efectividad del cribaje, la fiabilidad del cribaje y el tiempo de detección pueden ser modelizados más como una función del estado de la enfermedad que como su incidencia por defecto. Por ejemplo, en el caso de detección de cáncer de mama, se han desarrollado modelos matemáticos para explicar su evolución, y después se utilizan los resultados para diseñar el cribaje óptimo. La tasa de la progresión de la enfermedad fue explícitamente incluida como un factor que afecta la probabilidad de detección de la enfermedad. Otros modelos de detección de cáncer de mama incluyen dos atributos que conllevan información sobre la efectividad del cribaje: la mamografía y la edad del individuo. Modelizando estos factores como variables aleatorias se pueden derivar resultados analíticos para la sensibilidad (tasa positiva verdadera) y la especificidad (tasa negativa verdadera) de los procedimientos de test, utilizando repetidamente una aproximación basada en la estadística bayesiana. Este tipo de trabajos se ha extendido a otros tipos de cáncer, tales como los de colon y de cérvix, y también a enfermedades no-cancerígenas ni contagiosas, como puede ser la detección de glaucoma.

1.3.2. *Decisión clínica*

La contribución de la investigación operativa a la prevención de enfermedades y a la gestión de pacientes constituye una extensa y creciente área de investigación sobre servicios sanitarios. Este área se configura como un híbrido que reúne las matemáticas, el análisis estructural de la investigación operativa y las metodologías de obtención de soluciones (optimización, heurísticas, simulación) junto con un conocimiento profundo de los aspectos biológicos, económicos y sociales de la atención sanitaria.

El *análisis de decisiones* se ha convertido en un instrumento ampliamente utilizado a la hora de analizar decisiones médicas y la eficiencia de la práctica clínica. Uno de los primeros principios al estudiar una decisión en donde existe riesgo e incertidumbre es la determinación de los atributos, la estructura y los resultados de la decisión junto con sus probabilidades concomitantes. En general se usan árboles de decisión para mostrar las probabilidades, resultados y nodos de decisión para problemas complejos, tales como tratamientos de irradiación de tiroides durante la infancia o tratamientos de enfermedades en arterias coronarias. El aspecto principal a determinar en este tipo de árboles es la identificación de nodos críticos o de aquellas variables que son importantes a la hora de tomar la decisión o que el tomador de decisiones quiere modificar. En segundo lugar, se sitúa el análisis de sensibilidad de los parámetros del modelo, y con especial atención a las probabilidades fijadas *ex ante,* para examinar cambios en los resultados y la robustez del modelo.

Otras aplicaciones de decisión clínica se encuentran en lo que se denomina *mejora del resultado,* es decir, el intentar mejorar la precisión de los diagnósticos en condiciones agudas y/o crónicas, y en la mejora de los resultados de las estrategias de cribaje. Un ejemplo de este tipo de aplicaciones consiste en intentar modelizar el tratamiento farmacológico de la hipertensión. En primer lugar se desarrolla una representación estocástica del tratamiento de hipertensión que predice el resultado de regímenes de tratamiento anti-hipertensivo observados en situaciones clínicas. Esto se realizó en la práctica utilizando inputs de expertos para determinar los diferentes protocolos de tratamiento y los resultados asociados. Se incluyeron en el estudio la efectividad del régimen de tratamiento, efectos colaterales, costes, síntomas de hipertensión, y se utilizaron como resultado las diferencias entre la presión sanguínea presentada por el paciente y la normal. Con estos parámetros se pudieron determinar las probabilidades de que un paciente tuviera la presión sanguínea controlada en una visita determinada. A continuación se obtuvieron, a través de expertos en el tema, las probabilidades de que un paciente no vuelva a la consulta cuando su presión no está controlada, y de que los pacientes se pierdan. Estos parámetros se incluyen en un algoritmo que determina el tratamiento óptimo con la información obtenida. Dichos resultados fueron validados comparándolos a estudios sobre resultados de tratamientos alternativos.

El análisis coste-beneficio (ACB) y el análisis coste-efectividad (ACE) se encuadran en una tercera área en donde la investigación operativa ha contribuido sustancialmente al avance de la gestión médica. En primer lugar, la investigación operativa puede ser útil para identificar los costes sanitarios netos, que incluyen costes directos de atención médica, más los costes debido a efectos colaterales negativos del tratamiento, menos el coste de la morbididad, ya que estos últimos representan los ahorros de los que se beneficia el sector sanitario debido a la prevención de la enfermedad. En segundo lugar su contribución radica en ayudar a determinar la efectividad sanitaria neta, es decir, el aumento de años de vida debido al tratamiento, mas el aumento de años ganados gracias a la eliminación de la morbididad. Después se restan los años de vida perdidos debido a efectos colaterales. Estos años, en general, se ajustan por calidad de vida (QALYs [7]). De aquí se obtiene el ratio coste-efectividad (costes dividido por los efectos netos). El análisis coste-beneficio, por otro lado, pone un valor monetario a los efectos netos y después examina si los beneficios son superiores a los costes. Sin embargo, este sistema suele ser controvertido porque obliga a valorar la vida humana en algunos casos. Existen varias técnicas para obtener este valor, que incluyen la disponibilidad a pagar, o la valoración de capital humano. Por su parte, el ACE determina la eficiencia relativa de un tratamiento determinado, y comparando con tratamientos de enfermedades o condiciones similares se puede identificar una frontera eficiente y un tratamiento «óptimo». Sin embargo, esta comparación en muchos casos resulta difícil porque cada tratamiento puede tener estructuras diferentes tanto de proceso como de resultado. Por lo tanto, la mayoría de decisiones se destinan más a la mejora de la eficiencia que a la obtención de optimalidad.

Un ejemplo de aplicación de técnicas de investigación operativa dentro del ACB es la decisión de inversión en tecnología. Por ejemplo, se puede simular una función de utilización de equipamiento y, a través del análisis de Monte Carlo junto con la información de los costes de inversión, se puede determinar bajo qué patrones de utilización la inversión es rentable.

Como se ha visto en este capítulo, la investigación operativa se utiliza en muchos de los ámbitos relacionados con la atención sanitaria. Cada vez más se utilizan estas técnicas para ayudar al decisor a

[7] «Quality-Adjusted Life Years», años de vida ajustados por calidad.

tomar decisiones complejas en un mundo en donde los presupuestos son cada vez más restrictivos y los recursos más escasos. Por otro lado, gracias al avance en el mundo de la informática tanto a nivel de equipos («hardware») como de programas («software») específicos, su aplicación resulta cada vez más sencilla y su abanico cada vez más amplio. Por un lado, problemas que no podían resolverse hace una década hoy son triviales en términos de tiempo de ordenador; por otro lado, cada vez más se desarrollan sistemas «expertos» que simplifican y por lo tanto facilitan mucho la programación de los modelos. También existen muchas revistas especializadas y manuales de investigación operativa para todos los niveles.

1.4. Dos ejemplos de aplicación de la investigación operativa en el ámbito sanitario

1.4.1. *Planificación y asignación de recursos en un sistema de salud mental* [8]

El organismo responsable del sistema de salud mental de un país, región o ciudad, tiene, entre otras, la responsabilidad de planificar un programa de apoyo e integración de enfermos mentales de esa región, y gestionar los recursos y servicios de tratamiento para este grupo de enfermos. Este ha sido el tema principal del trabajo desarrollado por H. Stephen Leff, Maqbool Dada y Stephen C. Graves [1986] que se ofrece a continuación. En este estudio se presenta un modelo general para la representación del problema de planificación y asignación de recursos de un sistema de salud mental, basado en técnicas cuantitativas tales como el relativo a cadenas de Markov y a la programación lineal y lineal entera. Este modelo es utilizado como una herramienta de ayuda a la decisión para los responsables del sistema de salud mental, permitiendo hacer un uso más efectivo de los recursos, simular escenarios futuros y dar respuesta a preguntas del tipo «¿qué pasaría sí...?». Su implementación se ha realizado a través de los Sistemas de Apoyo Comunitario (Community Support Systems), responsables del Sistema de Salud Mental, en los Estados Unidos.

Las principales respuestas del modelo se dirigen a la planificación y asignación de recursos a lo largo del tiempo y a la asignación de servicios a categorías de enfermos, respetando la cantidad de recursos disponible y que cada enfermo reciba un tratamiento ade-

---

[8] H. Stephen Leff, Maqbool Dada y Stephen C. Graves [1986], «An LP planning model for a mental health community support system», *Management Science*, 32, n.º 2, 139-155.

cuado a su categoría. Además, el modelo permite hacer un seguimiento y evaluación del programa.

La construcción del modelo multi-período tiene tres fases:

1. Definir las categorías de enfermos. Se pretende obtener una clasificación de los enfermos en función de sus necesidades y de su respuesta a determinado tratamiento.

2. Definir un conjunto de servicios. Obtener una lista de servicios de acuerdo con las necesidades de los enfermos y con la disponibilidad de los recursos, basada en la experiencia y conocimientos médicos.

3. Planificar y asignar los recursos. El objetivo es asignar los conjuntos de servicios a las distintas categorías de enfermos a lo largo del tiempo, usando solamente los recursos disponibles en cada período y minimizando (o maximizando) un determinado objetivo.

La metodología usada en la primera fase se basa en técnicas estadísticas, para la recogida de datos y la determinación del historial del enfermo. Las categorías de los enfermos se definen en base a la experiencia y conocimientos médicos. A lo largo del tiempo, los enfermos pueden salir del sistema, nuevos enfermos pueden entrar, y también los enfermos pueden cambiar de categoría como respuesta positiva o negativa a un tratamiento. Las cadenas de Markov son una técnica estadística muy estudiada que permite la representación de estos cambios por medio de las probabilidades de transición. La segunda fase se hará con base a la experiencia y conocimientos médicos.

En la tercera fase, relativa a la planificación y asignación de recursos, la metodología usada se basa en técnicas de programación lineal y programación lineal entera. La técnica cuantitativa de programación lineal es una de las más usadas para la asignación óptima de recursos en una organización. El problema se formula como un modelo multi-período de programación lineal, definiendo la función objetivo de minimización (o maximización); por ejemplo, minimizar el número de enfermos en determinadas categorías al final del horizonte temporal, construyendo las restricciones relativas a la disponibilidad de los recursos y garantías de que todos los enfermos tengan tratamiento. El paso siguiente es la resolución del problema mediante un programa informático para la obtención de la solución óptima. También se pueden simular diferentes escenarios cambiando las restricciones y/o la función objetivo en el modelo multi-período.

En el estudio comentado, los autores citan que muchos responsables de sistemas de salud han usado este modelo con éxito en la toma de decisiones estratégicas y en la definición de políticas relacionadas con la planificación y asignación de recursos en un sistema de salud. Por ejemplo, con este tipo de modelos se pueden obtener distintos escenarios, variando el presupuesto y estudiar el impacto de estos cambios, sabiendo que para cada presupuesto se hace el mejor uso de los recursos disponibles, o analizar las consecuencias de abrir nuevos servicios. Con esta herramienta de ayuda a la decisión, las decisiones son basadas en la mejor asignación posible de los recursos disponibles, usando técnicas cuantitativas y simulación de escenarios, y no simples decisiones subjetivas.

1.4.2. *Programación de servicios de salud a domicilio* [9]

En la actualidad existen diversas organizaciones que ofrecen servicios de salud en el domicilio de los pacientes, tales como servicios de enfermería. El principal objetivo de estas organizaciones es hacer un uso eficiente de sus recursos para mejorar la calidad del servicio e incrementar la productividad, pero al mismo tiempo reduciendo costes. El principal recurso de estos servicios de salud es el personal de enfermería que se desplaza al domicilio de los pacientes. De este modo, uno de sus principales problemas es hacer la programación semanal y diaria de las visitas de cada enfermera disponible al domicilio de los pacientes y determinar el orden de las visitas, minimizando costes y garantizando una determinada calidad de servicio.

Begur, Miller & Weaver [1997] presentan un sistema de ayuda a la decisión para la programación semanal y diaria de las visitas de personal de enfermería a pacientes en su propio domicilio. El proyecto ha sido realizado por la Universidad de Alabama, EEUU, y por la «Visiting Nurses Association», que está usando el sistema. En los Estados Unidos existen más de 10.000 organizaciones que ofrecen servicios de enfermería o salud en general a domicilio, y siendo su tendencia la de crecer en el futuro próximo.

El sistema de ayuda a la decisión tiene los siguientes componentes: una base de datos, un sistema de información geográfico, un sistema de programación semanal y diario de visitas y un sistema de «interface» visual.

El sistema de bases de datos incorpora todos los datos relativos al personal, los enfermos, las visitas realizadas y a realizar, y un análi-

---

[9] S. V. Begur, D. M. Miller and J. R. Weaver [1997], «A Integral Spacial DSS for Scheduling and Routing Home-Health-Care Nurses», *Interfaces*, 27: 4, 35-48.

sis de productividad. En algunos casos, también se ha incorporado una conexión al sistema informático de contabilidad de la organización. La información obtenida en este sistema sirve de base para el sistema de programación y de interface visual.

Una de las características que más facilitan el uso de todo el sistema es la incorporación de un Sistema de Información Geográfica (SIG). El software escogido es el MAPINFO. Este sistema permite visualizar la programación del personal en global o en particular para cada categoría de profesionales, mediante mapas digitalizados de la región. Además, la modificación de los planes se torna muy sencilla y fácil de hacer.

El problema de programación semanal y diario de las visitas se resuelve mediante técnicas cuantitativas de optimización combinatoria, conocidas como heurísticas. Estas técnicas permiten obtener el orden de visitas y el horario para cada enfermero, minimizando los costes de viaje, o sea de trabajo no productivo, y garantizando que se respeta el horario de trabajo de los enfermeros y los requisitos especiales de cada paciente, en términos de tiempo y servicio de enfermería adecuado. La heurística implementada se basa en una adaptación al problema específico de la conocida heurística de Clark & Wright para problemas de ruteo de vehículos.

Finalmente, el sistema de interface visual permite al usuario ver la programación diaria en un mapa, y también otro tipo de información como, por ejemplo, el orden de las visitas a realizar en determinado día para cada enfermera/o, el horario de las visitas y su carga horaria total. También permite al usuario modificar la programación de una forma muy fácil y sencilla, y hacer análisis de distintos escenarios.

El sistema de ayuda a la decisión para la programación de las visitas ha sido adoptado por la «Visiting Nurses Association» y otras instituciones, y ha sustituido la programación manual, que era la herramienta más usada hasta ese momento. Para utilizar este sistema se necesita un PC, lo que permite hacer la programación en muy poco tiempo, de una forma consistente y fiable, y ahorrando tiempo de personal cualificado en la preparación de los planes. Las soluciones iniciales obtenidas por el sistema reducen de forma significativa los tiempos de viaje. El sistema de información geográfica y el interfase gráfico han permitido una fácil aceptación y aprendizaje del uso del sistema. Además, permite al usuario hacer análisis de diferentes escenarios, cambiando la solución inicialmente propuesta por el sistema. Otra prestación del sistema es la obtención de la documentación de cada enfermero, indicando para cada uno el orden de las visitas, horario y mapas con la ubicación de los domicilios a visitar, ahorrando de este modo la pérdida de tiempo en la búsqueda de los domicilios de los pacientes. Los autores calculan que para un escenario con 7 enfermeros y 40 visitas por día, se ahorra con este sistema cerca de 20.000 dólares al año en costes de viajes, de personal y preparación de la documentación.

2. PROGRAMACIÓN LINEAL I: FORMULACIÓN DE PROBLEMAS

2.1. Introducción

El desarrollo de la programación lineal, según muchos autores, ha representado uno de los avances científicos más importantes desde mediados del siglo XX. Actualmente es una herramienta utilizada en muchos campos de la administración, de la economía y de la ingeniería. Existen muchos libros de texto sobre el tema y miles de artículos científicos en revistas especializadas.

La programación lineal es un caso especial de la programación matemática, en donde todas las funciones que hay en el modelo son lineales: siempre tenemos una función objetivo lineal a optimizar (maximizar o minimizar), sujeta a restricciones lineales individuales. Las variables del modelo, que son continuas, únicamente pueden coger valores no negativos. Si bien puede parecer que estos supuestos quitan realismo al problema porque el modelador está limitado al uso de ecuaciones que quizás no son frecuentes en el mundo real, las técnicas de programación lineal se utilizan en un amplísimo espectro de problemas como, entre otros, de planificación y gestión de recursos humanos y materiales, de transporte, de planificación financiera y de organización de la producción. En definitiva, una extensa gama de problemas que aparecen en las áreas de tipo industrial, económico, administrativo, militar...

El término *programación* tiene su origen en la planificación de las actividades que se realizan en una organización tal como una fábrica, un hospital, una compañía aérea o un organismo público, en donde hay un objetivo a optimizar (maximización de beneficios, minimización de costes, maximización de la cobertura sanitaria, etcétera). No tenemos que confundir este término con la «programación» en referencia a la preparación de una serie de órdenes e instrucciones de un lenguaje informático en un ordenador.

Orígenes de la programación lineal

La programación lineal, si bien actualmente se utiliza frecuentemente para resolver problemas de decisión, era casi desconocida antes de 1947. Ninguna investigación significativa fue realizada antes de esta fecha, si bien hay que mencionar que, alrededor de 1823, el matemático francés Jean Baptiste Joseph Fourier parecía conocer el potencial del tema.

Un matemático ruso, Leonid Vitalievitx Kantorovitx, que publicó una extensa monografía en 1939, *Matematitxeskie Metodi Organisatsi i Planirovaniia Proisvodstva* (métodos matemáticos para la organización y planificación de la producción) fue el primer investigador en reconocer que una amplia gama de problemas de producción y distribución tenían una estructura matemática y, que por lo tanto, se puedan formular con un modelo matemático. Desgraciadamente sus propuestas fueron desconocidas tanto en Unión Soviética como en el occidente durante dos décadas. Durante este período, la programación lineal experimentó un gran desarrollo tanto en Estados Unidos como en Europa. Después de la segunda guerra mundial, funcionarios del gobierno americano consideraron que la coordinación de las energías de toda una nación debido al peligro de una guerra nuclear requeriría la utilización de técnicas científicas de planificación. Con la aparición del ordenador esto se hizo posible. Las Fuerzas Aéreas de los Estados Unidos se pusieron a trabajar intensamente. Se propuso un modelo de programación lineal por su simplicidad y aplicabilidad, sin dejar de dar un marco lo suficientemente amplio para representar actividades interdependientes que han de compartir recursos escasos. El sistema (como, por ejemplo, la producción industrial) se compone de diversas actividades relacionadas entre ellas (formación, fabricación, almacenaje, transporte, distribución y venta). Este fue el primer modelo de programación lineal conocido.

2.2. Formulación de modelos

En esta sección se presentan algunos ejemplos de los problemas con los cuales se puede encontrar una organización y cómo la programación lineal puede expresarlos matemáticamente.

2.2.1. *Un problema de asignación de personal*

El hospital Optsalud ha decidido ampliar su servicio de urgencias (abierto las veinticuatro horas) con la consiguiente necesidad de nuevo personal de enfermería. La gerencia del hospital ha estimado las necesidades mínimas de personal por tramos horarios para poder cubrir las urgencias que se presenten. Se definieron seis tramos de cuatro horas. La necesidad mínima de personal en cada tramo se indica en el Cuadro 2.1. Por otro lado, el departamento de recursos humanos ha informado a gerencia que los contratos laborales han de ser de ocho horas seguidas, según el Convenio firmado con los sindicatos, independientemente de los horarios de entrada y salida del personal. El problema es encontrar el número mínimo de personal necesario para cubrir la demanda.

CUADRO 2.1
Necesidades de personal por tramos horarios

	Tramos Horarios					
J	1 0:00-4:00	2 4:00-8:00	3 8:00-12:00	4 12:00-16:00	5 16:00-20:00	6 20:00-24:00
Personal N_j	5	7	13	12	4	3

Formulación del problema
En primer lugar, se tienen que definir las variables del modelo que queremos desarrollar. Como hemos de controlar un número de personal en cada turno, definimos X_j como la cantidad de personal que entra a trabajar en el turno j, en donde j = 1,....,6. Es decir, hay una variable para cada turno.

Las restricciones del modelo tienen que reflejar la necesidad de que la cantidad de personal que entren en el período j más el número de personas que entraron a trabajar en el turno j – 1 sean suficientes para cubrir las necesidades del turno j (N_j). Esta situación queda reflejada en el Cuadro 2.2. En esta tabla, un trabajador que entra a trabajar, por ejemplo, a las 4:00, trabajará en los turnos 2 y 3, y por tanto, contribuirá a cubrir las necesidades de estos dos turnos. En otras palabras, el turno j estará siendo atendido por X_{j-1} y X_j. En consecuencia, tendremos que $X_{j-1} + X_j$ (el personal que trabaja durante el turno j) tiene que ser, como mínimo, igual a N_j, que es el número mínimo de personal de enfermería necesario para este turno. En términos matemáticos, la restricción es la siguiente:

$$X_{j-1} + X_j \geq N_j$$

Habrá una restricción para cada horario de entrada.

CUADRO 2.2
Necesidades de personal por tramos horarios

	Tramos Horarios					
	1 0:00-4:00	2 4:00-8:00	3 8:00-12:00	4 12:00-16:00	5 16:00-20:00	6 20:00-24:00
0:00 4:00 8:00 12:00 16:00 20:00	X_1 X_6	X_1 X_2	X_2 X_3	X_3 X_4	X_4 X_5	X_5 X_6
Personal N_j	5	7	13	12	4	3

El objetivo de la gerencia consiste en la minimización del número total de personal de enfermería necesario para cubrir las necesidades diarias. Este número será igual a $X_1 + X_2 + X_3 + X_4 + X_5 + X_6$, que representa la suma del número de personal que entra en cada período. Finalmente, el modelo matemático es el siguiente:

$$\text{Min } Z = \sum_{j=1}^{6} X_j$$

Sujeto a:

$$\begin{aligned}
X_6 + X_1 &\geq 5 \\
X_1 + X_2 &\geq 7 \\
X_2 + X_3 &\geq 13 \\
X_3 + X_4 &\geq 12 \\
X_4 + X_5 &\geq 4 \\
X_5 + X_6 &\geq 3 \\
X_j \geq 0, \quad j &= 1, \ldots, 6
\end{aligned}$$

2.2.2. *Un problema de asignación de recursos*

El gerente del hospital Muchsalud ha observado que algunos de sus servicios tienen capacidad ociosa. Siguiendo una propuesta realizada por el equipo médico, esta capacidad ociosa podría aprovecharse para introducir dos tipos nuevos de cirugía, A y B. Tanto los pacientes de tipo A como los de tipo B tienen que pasar primero por una sala de pre-cirugía y, una vez pasado por el quirófano, tienen que estar en observación en una sala posoperatoria, que no existe de momento. El equipo médico ha estimado el tiempo medio que necesita cada paciente de tipo A y de tipo B en cada uno de los servicios pre-quirúrgico (PQ), quirúrgico (QI) y posoperatorio (PO). La experiencia en un hospital similar muestra que por cada tres pacientes de tipo A que llegan al hospital como mínimo llega uno de tipo B. Por otra parte, se ha estimado el coste de cada paciente en los diferentes servicios. El Cuadro 2.3 muestra los datos del problema, teniendo en cuenta que la capacidad ociosa es en horas mensuales y el coste en miles de pesetas.

CUADRO 2.3
Estimaciones horarias de las cirugías A y B

	Horas necesarias de cirugía		Capacidad ociosa
	A	B	
Sala PQ Sala QI Sala PO	2 1 4	3 2 2	150 160
Coste (€)	15	12	

Como el servicio posoperatorio (PO) aún no existe, el gerente argumenta que para justificar su creación tiene que utilizarse durante un mínimo de 135 horas al mes. Por otra parte, el presupuesto mensual asignado a las nuevas cirugías es de 1.000 euros. El gerente quiere saber cual será el número máximo de pacientes que podrán ser operados al mes.

Formulación matemática del problema
Primero definimos las variables del modelo. Sean X_1 y X_2 el número total de pacientes por mes que pueden ser tratados con la cirugía A y B, respectivamente. A continuación se presentan las restricciones.

Se ha establecido que en la sala PQ se dispone de 150 horas. En otras palabras, la utilización de esta sala no puede sobrepasar las 150 horas. Como cada uno de los pacientes de tipo A y de tipo B consumen 2 horas y 3 horas en esta sala respectivamente, el número total de horas mensuales consumidas en PQ para los dos tipos será igual a $2X_1 + 3X_2$. Este número tiene que ser inferior o igual a las 150 horas. La restricción será la siguiente:

$$2X_1 + 3X_2 \leq 150$$

El mismo razonamiento puede ser utilizado para determinar el número límite de horas en la sala QI. Como el total de horas consumidas será igual a $2X_1 + 3X_2$, y hay un máximo de 160 horas disponibles, la restricción sobre QI será:

$$X_1 + 2X_2 \leq 160$$

El gerente ha determinado que, para viabilizar los nuevos tratamientos, se tiene que ocupar la nueva sala PO durante un mínimo de 135 horas al mes. Como el número de horas mensuales que se utilizará en PO es igual a $4X_1 + 2X_2$, tendremos que:

$$4X_1 + 2X_2 \geq 135$$

La experiencia en otros hospitales muestra que, por cada 3 pacientes de tipo A, viene como mínimo un paciente de tipo B. Matemáticamente, esto se expresa como:

$$X_1 / 3 \leq X_2$$

que es equivalente a:

$$X_1 - 3X_2 \leq 0$$

Finalmente, el gasto mensual realizado en las dos cirugías no puede exceder el millón de pesetas. Como cada paciente de tipo A y de tipo B cuesta 15 euros y 12 euros, respectivamente, el gasto total mensual será de $15X_1 + 12X_2$, cantidad que no puede exceder el presupuesto, tendremos que:

$$15X_1 + 12X_2 \leq 1.000$$

Ahora se necesita formular el objetivo. El gerente quiere saber el número máximo de enfermos de tipo A y de tipo B que se puede atender cada mes. Simplemente, tendremos que si Z es este número, el objetivo se expresará como:

$$\text{Max } Z = X_1 + X_2$$

En resumen, la formulación del problema es la siguiente:

$$\begin{aligned} \text{Max } Z = \quad & X_1 + X_2 \\ \text{s. a.} \quad & \\ & 2X_1 + 3X_2 \leq 150 \\ & X_1 + 2X_2 \leq 160 \\ & 4X_1 + 2X_2 \geq 135 \\ & X_1 - 3X_2 \leq 0 \\ & 15X_1 + 12X_2 \leq 1.000 \\ & X_1, X_2 \geq 0 \end{aligned}$$

2.2.3. *Un problema de transporte*

El hospital Saludmuch pertenece a la Compañía de Seguros Todosalud, S.A. Esta sociedad tiene un Centro de Asistencia Primaria (CAP) en m pueblos y ciudades de una región (un CAP en cada centro urbano). Para obtener un buen funcionamiento global del servicio y poder planificar el número de visitas en función del personal previsto en cada CAP y de su dimensión, Todosalud, S.A., ha decidido organizar el servicio de tal forma que todos sus asegurados tengan un CAP de referencia asignado, pero que sea éste el más cercano posible a su lugar de residencia. En la región hay n ciudades y pueblos ($n > m$) y la compañía sabe cuántos asegurados tiene en cada uno de ellos. El objetivo es asignar a los asegurados a los CAPs minimizando el coste o la distancia total.

En primer lugar se definen los parámetros necesarios para formular el modelo. Sea a_j el número de asegurados en el centro urbano j, $j = 1,..., n$. Sea b_i el número total de asegurados que el CAP i puede tener asignados como máximo, $i = 1,...,m$. Se define c_{ij} como el coste de desplazamiento entre i y j.

Como se necesita conocer cuántas personas del centro urbano j serán asignadas al centro i, se define la variable X_{ij} como el número de personas que provienen del centro urbano j que serán atendidas por el CAP i.

Una vez definidos los parámetros y las variables, necesitamos definir las restricciones del modelo. En este problema hay dos tipos de restricciones. La primera viene definida por la capacidad de atención máxima de los CAPs. El número total de asegurados asignados al CAP i no puede exceder su capacidad b_i. En términos matemáticos:

$$\sum_{j=1}^{n} X_{ij} \leq b_i \qquad i=1,\ldots,m$$

El segundo grupo de restricciones tiene que considerar que hemos de asignar la totalidad de los asegurados de Todosalud SA de cada centro urbano j a los CAPs existentes.

$$\sum_{i=1}^{m} X_{ij} = a_j \qquad j=1,\ldots,n$$

Finalmente, se tiene que formular el objetivo de minimización total de la distancia o coste total del sistema. Este viene definido por:

$$c_{11}X_{11} + c_{12}X_{12} + \ldots + c_{1n}X_{1n} + \ldots + c_{ij}X_{ij} + \ldots + c_{m1}X_{m1} + \ldots + c_{mn}X_{mn}$$

que podemos re-escribir en forma compacta como:

$$\text{Min } Z = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}X_{ij}$$

En resumen, la formulación completa del modelo es la siguiente:

$$\begin{aligned}
&\text{Min } Z = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}X_{ij} && \\
&\sum_{j=1}^{n} X_{ij} \le b_i && i=1,\ldots,m \\
&\sum_{i=1}^{m} X_{ij} = a_j && j=1,\ldots,n \\
&X_{ij} \ge 0, && i=1,\ldots,m \quad j=1,\ldots,n
\end{aligned}$$

Se tiene que observar que este problema presenta una peculiaridad que no está en la formulación. Para que el problema tenga una solución factible, el número total de asegurados no puede exceder la capacidad total de los CAPs. Es decir, existe la siguiente restricción implícita en el modelo:

$$\sum_{i=1}^{m} b_i \ge \sum_{j=1}^{n} a_j$$

Si esto no se verificara, el problema no tendría solución.

2.2.4. *Un problema de programación financiera*

La compañía de seguros Todosalud, S.A., está preparando su plan de inversiones para los próximos dos años. Actualmente, la empresa tiene 2 millones de pesetas para invertir y espera ingresar, gracias a inversiones pasadas, un flujo de dinero al final de los meses, 6, 12 y 18 próximos. Por otra parte, la empresa quiere expandirse y tiene dos propuestas sobre la mesa. La primera es asociarse con la empresa Sanimas, S.A., y la segunda con la empresa

Buenavida, S.A. En el Cuadro 2.4 se muestra el flujo de caja de Todosalud, S.A., si entrara con un 100 % en cada uno de los proyectos.

CUADRO 2.4
Flujo de Caja de Todosalud, S.A.
(miles de pesetas)

	Inicial	6 meses	12 meses	18 meses	24 meses
Inversiones pasadas .		500	400	380	
Sanimas, S.A.	−1.000	−700	1.800	400	600
Buenavida, S.A..	−800	500	−200	−700	2.000

Debido al actual nivel de endeudamiento, a Todosalud, S.A., no se le permite pedir préstamos. Pero sí que puede, a cada seis meses, invertir sus fondos excedentes (es decir, aquellos que no ha invertido en ningún proyecto) en un fondo que le daría un 7 % cada seis meses. Por otro lado, Todosalud, S.A., puede participar en cada uno de los proyectos con un nivel inferior al 100 % y, consecuentemente, el flujo de caja se reducirá en la misma proporción. Es decir, que si decide entrar por ejemplo con el 50 % en el proyecto de Buenavida, el flujo correspondiente también se reducirá en la misma proporción. El problema que se plantea Todosalud, S.A., es cuánto invertir en cada proyecto para maximizar el dinero en efectivo que tendrá la empresa en dos años.

Formulación matemática del problema
Siguiendo nuestro esquema habitual, una vez el problema ha sido identificado y los parámetros del modelo han sido definidos, se tienen que definir las variables. Sea X_1 el porcentaje de participación en el proyecto Sanimas y X_2 el porcentaje de participación en el proyecto Buenavida, S.A. ($0 \leq X_1 \leq 1$, $0 \leq X_2 \leq 1$). Por otro lado, sean S_0, S_6, S_{12} y S_{18} el dinero que se depositará en el fondo en los períodos 0, 6 12 y 18, respectivamente.

Para formular las restricciones del modelo se utilizará un razonamiento secuencial. La empresa dispone de 2 millones de pesetas hoy (período 0) y las quiere gastar considerando las opciones siguientes:

1. Participar en el proyecto Sanimas, que implicaría desembolsar 1.000.000 X_1 pesetas en el período 0;
2. Participar en el proyecto Bonavida, teniendo que gastar 800.000 X_2;

3. Depositar el dinero al 7 %

Estas opciones no son excluyentes entre ellas. Por lo tanto, se tiene que cumplir la siguiente ecuación de equilibrio:

$$2.000 = 1.000X_1 + 800X_2 + S_0$$

Al cabo de seis meses, la empresa ingresará 500.000 ptas. gracias a inversiones realizadas anteriormente. También el dinero depositado en el fondo en el período anterior estará a disposición junto con los intereses: $S_0 + 0{,}07S_0$. Por otra parte, el proyecto Buenavida dará una entrada de dinero igual a $500.000X_2$. Con este dinero tendrá que hacer frente al compromiso adquirido con Sanimas, $700.000X_1$, y depositar lo que quede al 7 % una vez más. Matemáticamente:

$$500 + 500X_2 + 1{,}07S_0 = 700X_1 + S_6$$

En el período 12, la empresa recibirá 400.000 ptas. de inversiones anteriores, $1.800.000X_1$ ptas. del proyecto Sanimas y el dinero del fondo junto con los intereses. Con estos ingresos tendrá que cubrir el compromiso del proyecto Buenavida, $200.000X_2$ y depositar S_{12} ptas. en el fondo. En términos matemáticos:

$$400 + 1.800X_1 + 1{,}07S_6 = 200X_2 + S_{12}$$

En el período 18, los ingresos que tendrá la empresa vendrán de inversiones anteriores (380.000 ptas.), del proyecto Sanimas ($400.000X_1$) y del depósito realizado en el período anterior incluyendo los intereses ($1{,}07S_{12}$). Con este dinero tendrá que realizar un gasto de $700.000X_2$ en el proyecto Buenavida y el resto puede volver a ponerlo en el fondo (S_{18}). Es decir:

$$380 + 400X_1 + 1{,}07S_{12} = 700X_2 + S_{18}$$

Finalmente, al cabo de dos años (período 24), la empresa tendrá únicamente ingresos y no tendrá ningún gasto. Los ingresos provienen de los dos proyectos ($600.000X_1 + 2.000.000X_2$) y del dinero depositado en el período anterior, $1{,}07S_{18}$. Si se define Z como los ingresos realizados en el período 24 en miles de pesetas, tendremos que:

$$Z = 600X_1 + 2.000X_2 + 1{,}07S_{18}$$

que no es más que el objetivo del problema: Maximizar los ingresos al cabo de dos años.

Finalmente, como sólo se puede invertir un máximo de 100 % en cada proyecto, las variables X_1 y X_2 no pueden exceder la unidad. Por lo tanto, hay que añadir las restricciones siguientes:

$$X_1 \leq 1$$

$$X_2 \leq 1$$

En resumen, reordenando los términos tendremos que el programa lineal se escribe de la forma siguiente:

$$\text{Max } Z = 600X_1 + 2.000X_2 + 1{,}07S_{18}$$

$$\begin{array}{rrrrrrl}
\text{s. a} & & & & & & \\
1.000X_1 & 800X_2 & +S_0 & & & & = 2.000 \\
700X_1 & -500X_2 & -1{,}07S_0 & +S_6 & & & = 500 \\
-1.800X_1 & 200X_2 & & -1{,}07S_6 & +S_{12} & & = 400 \\
-400X_1 & 700X_2 & & & -1{,}07S_{12} & +S_{18} & = 380 \\
X_1 & & & & & & \leq 1 \\
 & X_2 & & & & & \leq 1
\end{array}$$

$$X_1, X_2, S_0, S_6, S_{12}, S_{18} \geq 0$$

2.3. Aplicación real: fabricación de válvulas cardiacas [10]

En el laboratorio American Edwards fabricaron en 1981 una nueva válvula cardiaca biológica para ser utilizada en seres humanos. Las válvulas se fabricaban utilizando corazones de cerdo comprados a varios distribuidores. Como no siempre las válvulas cardiacas de cerdo compradas tenían el mismo tamaño que las válvulas humanas, la empresa tenía dificultades para mantener un *stock* suficiente para poder satisfacer la demanda (a veces recibía un envío de corazones de cerdo que era inservible). Sid Hilal y Warren Erikson desarrollaron un programa lineal para seleccionar la combinación de proveedores que se acercaría más a la medida correcta de las válvulas cardiacas humanas. Resultado: una reducción de stock valorada en 1,9 millones de dólares y ahorros anuales de 1,5 millones de dólares.

En el laboratorio American Edwards, la disponibilidad del producto era de una importancia capital y la política de la empresa descarta-

---

[10] S. Hilal y W. Erikson (1981): «Matching supplies to save lives: Linear Programming and the Prodution of Heart Valves», *Interfaces*, 11(6): 48-56.

ba el uso de cualquier método tradicional de gestión de stock para controlar el margen de seguridad de la válvulas cardiacas almacenadas. El objetivo de la empresa era siempre tener el volumen de demanda de seis a doce meses en *stock* de seguridad. Conseguir este objetivo era difícil porque no podían comprar las válvulas con medidas específicas. Se desconocía la medida de una válvula dentro de una carga hasta que no se procesaba todo el envío en el laboratorio. El resultado es que a veces la carga era inservible.

El estudio realizado por Hilal y Erikson mostró que la mayoría de los distribuidores entregaban los corazones de cerdo con una distribución del tamaño de las válvulas bastante estables. Estas distribuciones se utilizaron dentro de un programa lineal para seleccionar la mejor combinación de distribuidores. La mejor manera de explicar el modelo es utilizando un ejemplo simple. Supongamos que la compañía compra a dos proveedores A y B y en su pedido se pueden encontrar tres tamaños de válvulas 1, 2 y 3. Los datos históricos muestran que los envíos del proveedor A tienen un 30 % de medida 1, un 50 % de medida 2 y un 20 % de medida 3. La distribución del proveedor B es 10 %, 60 % y 30 %, respectivamente. Supongamos que los costes totales de compra y manipulación para las medidas 1, 2 y 3 son 10, 14 y 12 dólares, respectivamente.

El valor esperado del coste de una válvula del proveedor A es $0{,}3(10) + 0{,}5(14) + 0{,}2(12) = 12{,}4$. El coste esperado para el proveedor B es $0{,}1(10) + 0{,}6(14) + 0{,}3(12) = 13$. El objetivo del programa lineal es:

$$\text{Min } Z = 12{,}4A + 13B$$

en donde A y B representan la cantidad de corazones comprados a cada proveedor. Si la demanda para los tres tamaños es igual a 100, 300 y 250 unidades, las restricciones serán:

$$0{,}3A + 0{,}1B \leq 100$$

$$0{,}5A + 0{,}6B \leq 300$$

$$0{,}2A + 0{,}3B \geq 250$$

El modelo real implementado por los autores tenía más de 20 proveedores y 30 medidas de válvulas. Pero la estructura del modelo era básicamente igual a la de nuestro ejemplo. El objetivo era la minimización de los costes de compra de los cerdos, con restricciones en el número de corazones de cada medida que se tenían que comprar.

El programa lineal permitió que la compañía cumpliese la demanda comprando menos corazones que antes. El modelo también sirvió para fijar precios, programar la producción y analizar el diseño de las válvulas nuevas.

2.4. Problemas

2.1. Una compañía de seguros sanitarios ha decidido atender dos nuevos tipos de pacientes en sus ambulatorios A, B y C, que tienen capacidad sobrante. Las previsiones indican que pueden venir un total de 100 pacientes de tipo 1 y 150 de tipo 2. Estos pacientes pueden ser atendidos en cualquier ambulatorio, excepto en el ambulatorio A, en donde no pueden atender a pacientes de tipo 2 por falta de equipos adecuados. La empresa quiere saber a cuántos pacientes podrá atender en cada uno de los ambulatorios para minimizar los costes totales de atención. Los costes de atención por paciente y ambulatorio se indican en la tabla siguiente:

Ambulatorio	Coste por paciente		Capacidad ociosa (ambos pac.)
	Pacientes 1	Pacientes 2	
A	26	—	80
B	28	33	50
C	24	28	120

Definir las variables y formular el problema.

2.2. El ministerio de sanidad decide hacer una campaña anti-tabaco mediante anuncios en la radio y la televisión. Su presupuesto limita los gastos de publicidad a 1.000.000 de ptas. por mes. Cada minuto de anuncio en la radio cuesta 5.000 ptas. y cada minuto en la televisión cuesta 100.000 ptas. El ministerio desearía utilizar la radio cuando menos dos veces más que la televisión. La experiencia pasada muestra que cada minuto de publicidad por televisión generará en términos generales 25 veces más impacto que cada minuto de publicidad por la radio. Formular el problema que determine la asignación óptima del presupuesto mensual para anuncios en radio y televisión que maximiza el impacto total.

2.3. La clínica Coratac ofrece cuatro tipos de servicios: cirugía plástica, dermatología, cirugía ortopédica y neurocirugía. Después de examinar los archivos contables se ha calculado que cada paciente, en cada una de las especialidades, contribuye al beneficio de la clínica de la manera siguiente: plástica, 100; dermatología, 200; ortopedia, 150, y neurocirugía, 180. Los médicos están convencidos de que el número de pacientes de cada especialidad que se atienden semanalmente no es el adecuado. La clínica quiere saber cuál sería el volumen semanal óptimo de pacientes en cada especialidad teniendo en cuenta los recursos de la clínica.

Especialidad	Horas necesarias por paciente				
	Laborat.	Rayos X	Terapia	Cirugía	Médicos
Plástica	5	2	1	4	10
Dermatología	5	8	10	8	14
Ortopedia	2	1	0	16	8
Neurocirugía	4	5	8	10	12
Horas disponibles semanales	200	140	110	240	320

La clínica no tiene ningún problema con la demanda de sus servicios y tiene acceso a tantos pacientes como quiera. Adicionalmente, ha decidido que limitará sus servicios de plástica y ortopedia combinados a un máximo de 120 pacientes. Formular el problema.

2.4. El laboratorio MacAsp prepara dos tipos de medicinas para el dolor de cabeza: el fantástico Resacón y el magnífico Jaquecón. Los dos se venden en forma de jarabe en frascos de 100 mg y se obtienen mezclando ácido acetil-salicílico (A) y paracetamol (P). El laboratorio se permite una cierta flexibilidad en las fórmulas de estos productos. De hecho, las restricciones son: (1) el Resacón tiene que tener un máximo de 75 % de A; (2) el Jaquecón ha de tener un mínimo de 25 % de A y un mínimo de 50 % de P. El departamento comercial piensa poder vender como máximo 400 frascos de Resacón y 300 de Jaquecón. El precio de venta es de 150 ptas. por un frasco de Resacón y de 200 ptas. por uno de Jaquecón. Los costes de los componentes son: 80 ptas. Por 100 mg de A y 120 ptas. por 100 mg de P. El laboratorio quiere maximizar el ingreso neto por venta. Formular el problema.

2.5. La red de hospitales Salutmolt ha detectado que podría atender a más pacientes y ha decidido ampliar sus servicios con tres nuevas especialidades X, Y y Z en tres hospitales A, B y C. Los beneficios esperados por paciente en cada especialidad X, Y y Z son 420, 360 y 300 euros, respectivamente. Los hospitales tienen capacidad de recursos humanos para atender 700, 800 y 450 pacientes independientemente de la especialidad. Un problema grave para el hospital es el número de horas disponibles de quirófano. Actualmente, se ha calculado que la capacidad ociosa es de 13, 12 y 5 horas de quirófano por día en A, B y C y cada especialidad necesita un máximo de 2, 1,5 y 1,2 horas por las cirugías X, Y y Z, respectivamente.

Para poder mantener una carga equilibrada de atención entre los diferentes hospitales, la gerencia ha decidido que la cantidad de pacientes asignados a cada hospital utilice el mismo porcentaje de capacidad adicional disponible. El gerente quiere saber cuántos pacientes podrán ser atendidos en cada hospital.

Formular el problema de decisión.

2.6. El departamento de finanzas de la mutua Hospimás ha decidido invertir 10 millones en fondos de inversión y ha considerado 20 fondos diferentes, cada uno de ellos con tipo de interés anual esperado diferente r_i, $i = 1,\dots,20$. Es decir, si la mutua pone x ptas., recibirá al cabo de un año $x\,(1 + r_i)$. Para mantener un equilibrio en la inversión y diversificar el riesgo, la mutua ha adoptado las siguientes reglas:

1. No invertir más de dos millones en un fondo único.

2. Si pone dinero en un fondo, lo tiene que hacer con un mínimo de 0,5 millones de ptas.

El objetivo es la maximización del rendimiento anual esperado. Formular el problema definiendo las variables de decisión.

2.7. Una empresa farmacéutica tiene m laboratorios. Todos los laboratorios producen el mismo medicamento. La gerencia quiere planificar la producción para los próximos T trimestres. En cada período t, cada laboratorio tiene una capacidad de producción igual a l_i^t, $i = 1,\dots,m$. La compañía envía el medicamento a n almacenes de distribución. Para poder atender la demanda, los requisitos mínimos de *stock* en cada almacén j y en cada período t son iguales a r_j^t. Para cada laboratorio, almacén y período, la empresa quiere determinar las variables siguientes:

- X_i^t = cantidad producida en i y enviada a los almacenes en el período t.

- S_i^t = cantidad producida en i y almacenada en el propio laboratorio en el período t.

- Z_{ij}^t = cantidad transportada del laboratorio i al almacén j en el período t.

- W_j^t = cantidad almacenada en el almacén j al final del período t.

La empresa quiere que, para la suma de todos los períodos, la producción total conjunta de los laboratorios 1, 2 y 4 no sea superior a la producción conjunta de los laboratorios 3, 5 y 6. A causa del límite en la capacidad de *stock* de los laboratorios, no se pueden almacenar más de s_i medicamentos en cada período. Los costes de producción unitarios de los medicamentos son iguales a p_i. Los costes de almacenaje de un medicamento en el laboratorio i al final del período t es k_i^t y el coste de almacenaje al final del período t es h_j^t. El coste de transporte por medicamento es igual a c_{ij} entre el laboratorio i y el almacén j (suponer que el tiempo de transporte es virtualmente 0 comparado con la duración de un trimestre o, en otras palabras, lo que sale en el período t llega siempre en el mismo período t).

Diseñar un programa de planificación de la producción y *stock* de las fábricas y de *stocks* de almacenes que minimice el coste total de producción y transporte.

3. PROGRAMACIÓN LINEAL II: MÉTODOS DE RESOLUCIÓN

Hasta ahora hemos formulado matemáticamente algunos problemas de gestión y administración de recursos y de dinero. Pero un modelo matemático de decisión, por muy bien formulado que esté, no sirve de nada si no podemos encontrar una solución satisfactoria. Una de las características de la programación lineal es que, gracias a sus propiedades matemáticas, se consigue la solución óptima sin muchas dificultades. En esta sección examinaremos en primer lugar el método gráfico, sistema limitado a problemas con dos variables, y a continuación el método Simplex, el algoritmo más común para solucionar problemas lineales con muchas variables y restricciones.

3.1. El método gráfico

Este método es muy simple de utilizar, pero sólo puede ser aplicado a problemas con dos variables. Por otro lado, es muy útil para entender las propiedades matemáticas de la programación lineal. Consideremos el problema lineal siguiente:

$$\text{Max } Z = X_1 + 1{,}4X_2$$

$$\text{s. a.}$$

$$\begin{aligned} X_1 + 0{,}5X_2 &\leq 6 \\ 0{,}5X_1 + X_2 &\leq 6 \\ X_1 + X_2 &\leq 7 \\ 1{,}4X_1 + X_2 &\leq 9 \\ X_1, X_2 &\geq 0 \end{aligned}$$

En primer lugar, se dibuja en un gráfico cartesiano las restricciones del modelo, pero con signo de igualdad. Como se puede observar en la figura 3.1, la recta $X_1 + X_2 = 7$ separa el plano en dos semiplanos. Los puntos que corresponden al semiplano S_1 cumplen la restricción $X_1 + X_2 \leq 7$. Es decir, este semiplano contiene todas las combinaciones de X_1 y X_2 que satisfacen la restricción. Si dibujamos todas las restricciones y sus semiplanos correspondientes encontraremos que la región que forma la intersección de todos los semiplanos incluye todas las combinaciones de X_1 y X_2 que satisfacen todas las restricciones del modelo. Esta región se presenta en la figura 3.2 (el área entre los puntos A, B, C, D y E) y se conoce como la *región factible* o *espacio de soluciones* y es un conjunto convexo [11]. Cualquier problema de optimización con restricciones lineales tiene una región factible convexa. Cualquier solución de la región factible es conocida como *Solución factible*. Si la región está vacía no existen soluciones factibles (véase el ejemplo de la figura 3.4).

Ahora se tiene que escoger la solución factible que optimice nuestra función objetivo, que es $Z = X_1 + 1{,}4X_2$. Obsérvese que normalmente existen infinitas soluciones factibles y será la función objetivo quien escoja aquella que optimiza su valor. En la programación lineal la función objetivo también tiene forma lineal. Se trata de determinar el valor máximo de Z que cumpla todas las restricciones o, en otras palabras, encontrar los valores de X_1 y X_2, puntos dentro de la región factible, que maximicen Z.

La mecánica para lograr encontrar el punto óptimo se basa en la linealidad del objetivo. En este ejemplo, el objetivo se puede re-escribir de la forma siguiente:

$$X_2 = -(1/1{,}4)X_1 + (1/1{,}4)Z$$

Obsérvese que a medida que Z aumenta, la recta se desplaza paralelamente hacia fuera, ya que la pendiente es constante (en este caso igual a $-1/1{,}4$). Se trata de encontrar el valor de Z máximo, pero con la condición de que tiene que haber como mínimo un punto de la recta que atraviese la región factible. En el gráfico 3.2 esta recta se presenta para los valores de $Z = 7$, $Z = 9$ y $Z = 12.6$. La solución óptima del problema es $Z = 9$, que corresponde al punto $X_1 = 2$ y $X_2 = 5$. Para cualquier valor de Z superior a 9, no existirá ninguna solución factible. Para valores de Z inferiores a 9, existen muchas soluciones factibles, pero ninguna de ellas es óptima. Intuitivamente se puede ver que la

[11] Para cualquier pareja de puntos dentro del espacio factible, el segmento de línea que los une también se encuentra dentro del conjunto.

FIGURA 3.1
Visualización de una restricción

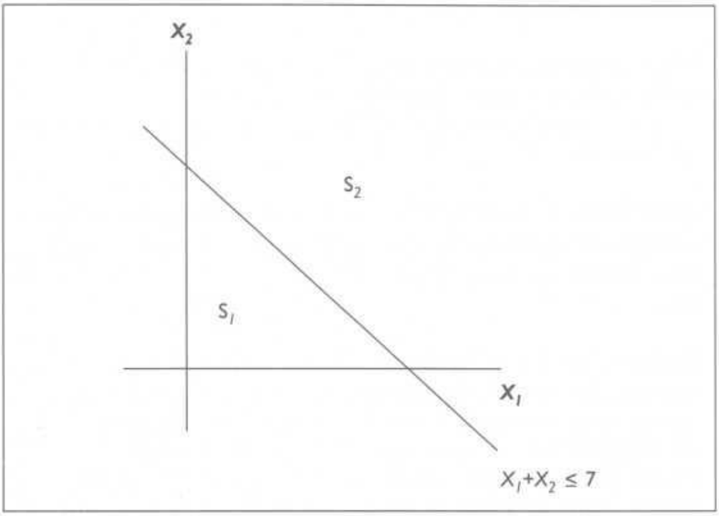
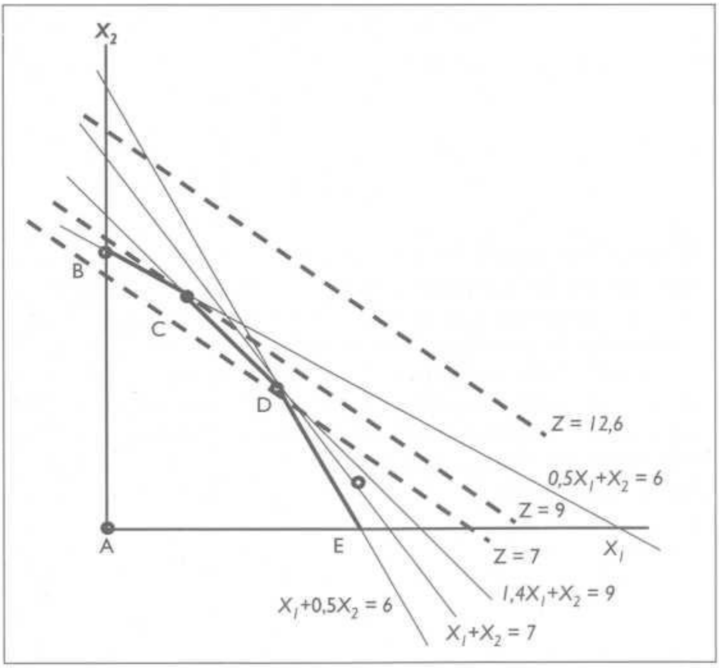


FIGURA 3.2
Solución gráfica del ejemplo



solución óptima siempre se producirá en un *punto extremo* o *vértice*, que en el gráfico no es más que el punto de intersección de dos o más restricciones.

Más formalmente, un punto de un conjunto convexo es un punto extremo si no hay ningún par de puntos del conjunto convexo en donde el segmento de línea que los une pase por el punto en cuestión.

Otra forma de obtener el óptimo es calcular el valor del objetivo en cada uno de los puntos extremos y escoger aquel punto extremo que da el mejor valor. Este punto dará el valor óptimo.

Existen situaciones en donde no hay una única solución, sino que pueden haber infinitas soluciones, o por el contrario, no existir solución alguna. Examinemos el primer caso con la ayuda de la figura 3.3. La recta correspondiente al objetivo tiene la misma pendiente que una de las restricciones. Es decir, que todas las combinaciones de las dos variables entre los puntos A y B cumplen las restricciones y maximizan el beneficio. Por otro lado, la figura 3.4 muestra una situación en donde no hay soluciones. La representación gráfica corresponde al programa lineal siguiente:

$$
\begin{aligned}
\text{Max } Z = \; & X_1 + 2X_2 \\
\text{s. a.} \; & \\
& X_1 \leq 6 \\
& X_1 \geq 8 \\
& X_1, X_2 \geq 0
\end{aligned}
$$

El método gráfico es sencillo de aplicar para encontrar la solución óptima de un programa lineal de optimización, pero únicamente cuando éste solo tiene dos variables de decisión. Desgraciadamente, la gran mayoría de problemas lineales aplicados tienen muchas más variables (algunos llegan a tener millones de ellas) y por lo tanto se hace inviable su utilización. En la sección siguiente se desarrolla un método bastante eficiente para encontrar soluciones óptimas de programas lineales con muchas variables y restricciones.

FIGURA 3.3
Infinitas Soluciones

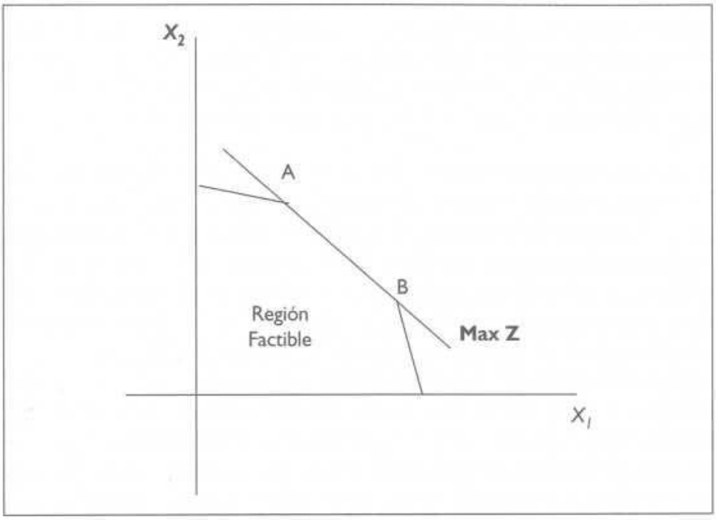
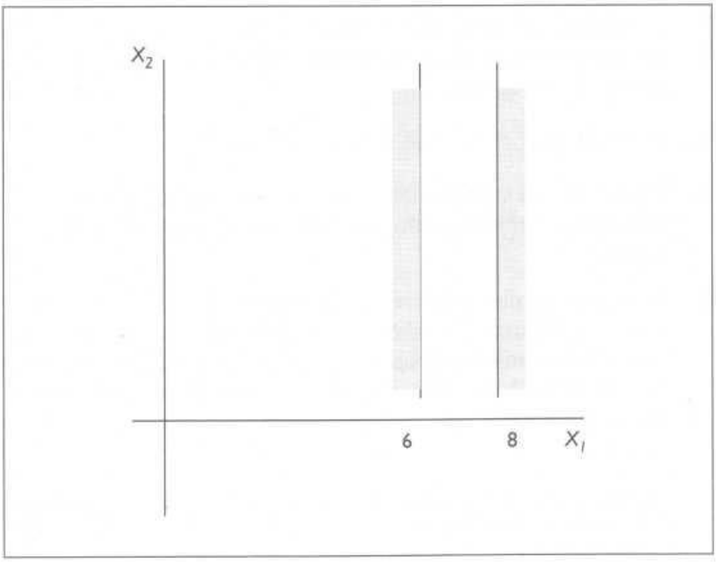


FIGURA 3.4
Inexistencia de soluciones



3.2. El método Simplex

El primer método formal para encontrar soluciones óptimas –el método Simplex– fue desarrollado por Dantzig en 1947 y mejorado por Charnes entre 1948 y 1952. Actualmente es el método más utilizado en la búsqueda de soluciones óptimas de programas lineales. En este apartado se examina su funcionamiento de forma simple e intuitiva.

En primer lugar recordemos cómo encontrábamos soluciones con el método gráfico. Primero formábamos un conjunto convexo con las restricciones del modelo. Segundo, se dibujaba la función objetivo fuera del conjunto convexo, dando un valor arbitrario al propio objetivo y se iba desplazando ésta paralelamente (ya que su pendiente es siempre constante) hasta encontrarse con un punto extremo. Intuitivamente, podemos ver que sea cual sea la función objetivo lineal, la solución óptima se encontrará en un punto extremo, como mínimo [12]. Esto reduce bastante el espectro de soluciones del problema, reduciendo la búsqueda del óptimo a los puntos extremos. Aún así, puede haber muchísimos puntos extremos en un problema. Por ejemplo, un problema grande con 2000 variables y 4000 restricciones tiene exactamente 2^{2000} puntos extremos, es decir, aproximadamente 10^{600}. Por lo tanto, tenemos que encontrar un método para reducir el número de soluciones factibles posibles de ser óptimas. Dantzig hizo estas mismas suposiciones (o eso creemos) y observó primero las características matemáticas siguientes:

1. El conjunto formado por las restricciones es convexo.
2. La solución siempre ocurre en un punto extremo.
3. Un punto extremo siempre tiene como mínimo dos puntos extremos adyacentes [13].

Y a partir de ellas desarrolló el método siguiente:

1. Encontrar una solución inicial factible en uno de los puntos extremos del conjunto convexo y calcular el valor de la función objetivo.

2. Examinar un punto extremo adyacente al encontrado en la etapa 1 y calcular el nuevo valor de la función objetivo. Si este nuevo valor mejora el objetivo, guardar la nueva solución y repetir la etapa 2. En caso contrario, ignorar la solución nueva y volver a examinar otro punto extremo.

---

[12] Decimos como mínimo, porque como hemos visto pueden existir (raramente) situaciones en donde hay más de una solución óptima; aún así, siempre habrá un punto extremo que dé el valor óptimo.

[13] Un punto extremo *A* es adyacente a un punto extremo *B* si no existe ningún punto extremo entre ellos. Por ejemplo, en la figura 3.2. los puntos *B* y *D* son adyacentes al punto *C*.

3. Regla de parada: cuando no existe ningún extremo adyacente que mejore la solución, nos hallamos en el óptimo.

Es decir, que vamos de punto extremo a punto extremo adyacente siempre que podamos mejorar la solución, hasta llegar a un punto en donde no existe ningún punto extremo adyacente al que nos encontramos. Esta solución es la óptima. Observemos de nuevo problema lineal presentado en la sección 3.1 y su correspondiente solución gráfica presentada en la Figura 3.2. Para encontrar una solución inicial en un punto extremo podemos fijar $X_1 = 0$ y $X_2 = 0$ y el valor del objetivo Z será igual a 0, solución que corresponde al punto extremo A en la Figura 3.2. Ahora examinamos el punto extremo adyacente B, que corresponde a los puntos $X_1 = 0$ y $X_2 = 6$ y $Z = 8{,}4$. Como el objetivo ha mejorado, mantenemos esta solución y volvemos a examinar los puntos extremos adyacentes a B. Como el punto A ya lo hemos visitado (y era claramente inferior), nos queda por ver el punto C. En este punto extremo $X_1 = 2$ y $X_2 = 5$ y $Z = 9$. De nuevo la solución ha mejorado y la guardamos como la mejor hasta ahora encontrada. Finalmente, D es el único punto extremo que nos queda por examinar y como en este punto $X_1 = 0$ y $X_2 = 3$ y $Z = 7{,}8$ el algoritmo se para y estamos en el óptimo.

Dantzig y más tarde Charnes desarrollaron un método matemático para poder efectuar estas operaciones, es decir, encontrar los valores de los puntos extremos adyacentes. Para poder ver cómo funciona, es necesario realizar las consideraciones siguientes:

Como hemos visto, un programa lineal está compuesto por una función objetivo que queremos optimizar (maximizar o minimizar), unas variables que denominaremos *estructurales* y un conjunto de restricciones. En general, podemos encontrar tres tipos de restricciones en función de la dirección de la desigualdad: $\geq$, $\leq$ o $=$. Toda restricción con los sentidos $\geq$ o $\leq$ pueden transformarse en una restricción con igualdad añadiendo una variable. Si la desigualdad tiene la dirección $\leq$, podemos añadir una *variable de holgura*. Por ejemplo, la restricción $X_1 + X_2 \leq 7$ se puede transformar en $X_1 + X_2 + X_3 = 7$. Si en la solución final del modelo la restricción se cumple con igualdad dados unos valores finales de X_1 y X_2 entonces la variable de holgura asociada a la restricción es igual a 0. Así mismo, si la restricción tiene la dirección $\geq$, podemos añadir una *variable de exceso* para obtener una ecuación lineal. Por ejemplo, la restricción $X_1 + X_2 \geq 12$ puede transformarse en $X_1 + X_2 - X_3 = 12$. La interpretación es la misma que en el caso anterior: si en la solución final $X_3 = 0$, la restricción se cumplirá con igualdad.

Con estas consideraciones, cualquier programa lineal con restricciones de desigualdad puede transformarse en un problema lineal con todas las restricciones con forma de igualdad sin alterar la naturaleza matemática del problema. Esta transformación se denomina *la forma canónica* o *forma aumentada* de un programa lineal. Si tenemos n variables y m restricciones con desigualdad, cuando escribimos la forma canónica del problema lineal tendremos m nuevas variables de holgura o exceso, es decir, un total de m + n variables y m restricciones. En resumen, tendremos que el conjunto de restricciones forma un conjunto de ecuaciones lineales con más variables que ecuaciones. En este caso existen infinitas soluciones del sistema y nuestro objetivo es escoger entre ellas la que optimice el valor de la función objetivo. Por otro lado, si tenemos un programa lineal con n variables, m restricciones con desigualdad y r restricciones con desigualdad, tendremos m + n variables y m + r restricciones con igualdad en la forma canónica. En este caso, para que el problema sea factible, se tiene que cumplir lo siguiente: $m + n \geq m + r$, el número de restricciones no puede superar el número de variables.

Si utilizamos el ejemplo de la sección 3.1, la forma canónica del problema será la siguiente:

$$\text{Max } Z = X_1 + 1{,}4X_2$$

s. a.

$$\begin{aligned}
X_1 + 0{,}5X_2 + X_3 &= 6 \\
0{,}5X_1 + X_2 + X_4 &= 6 \\
X_1 + X_2 + X_5 &= 7 \\
1{,}4X_1 + X_2 + X_6 &= 9
\end{aligned}$$

$$X_1, X_2, X_3, X_4, X_5, X_6 \geq 0$$

Los valores de las variables en los puntos extremos se presentan en el Cuadro 3.1.

CUADRO 3.1
Puntos extremos del ejemplo

	X_1	X_2	X_3	X_4	X_5	X_6	Número de variables positivas	Valor del objetivo
A	0	0	6	6	7	9	4	0
B	0	6	3	0	1	3	4	8,4
C	2	5	1,5	0	0	1,2	4	9
D	5	2	0	1,5	0	0	3	7,8
E	6	0	0	3	1	0,6	4	6
A	0	0	6	6	7	9	4	0

Diremos que una *solución aumentada* es una solución de la forma canónica del programa lineal. Una *solución básica* es una solución aumentada en un punto extremo. Una *solución básica factible* es una solución factible en un punto extremo. En nuestro ejemplo, los puntos A, B, C, D y E son soluciones básicas factibles.

A continuación examinaremos las propiedades algebraicas de las soluciones básicas. Obsérvese que en nuestro ejemplo tenemos dos variables estructurales X_1 y X_2 y cuatro variables de holgura X_3, X_4, X_5 y X_6, que suman un total de seis variables, y cuatro restricciones con igualdad o ecuaciones. Tenemos por lo tanto dos grados de libertad para encontrar soluciones. Para encontrar una solución determinada tenemos que fijar *a priori* dos variables para obtener entonces un sistema con cuatro variables y cuatro ecuaciones, que tendrá una solución única. En el método Simplex siempre se fija el valor de dos variables en 0. Estas variables se denominan *variables no básicas* y las restantes, *variables básicas.* La solución de este sistema de ecuaciones es una *solución básica.* Si todas las variables básicas son no-negativas, tenemos una *solución básica factible.* En el ejemplo tenemos que en cualquier punto extremo factible siempre tendremos dos variables iguales a 0 y 4 no-negativas. La explicación intuitiva de esta situación es la siguiente: si observamos un punto extremo en la Figura 2.1 veremos que en él pasan dos rectas correspondientes a dos restricciones con signo igual. Por lo tanto, dos variables de holgura asociadas a estas restricciones son iguales a 0. Estas son nuestras variables no-básicas. Si miramos el Cuadro 3.1, veremos que en cada punto extremo siempre hay cuatro variables positivas y dos con valor 0, excepto en el punto extremo D, en donde hay únicamente tres variables mayores que 0 (hay que recordar que las condiciones de no-negatividad $X_1, X_2 \geq 0$ también son restricciones del problema). Simplemente, esto es debido a que por este punto pasan tres restricciones.

El Cuadro 3.1 nos puede ayudar a entender cómo funciona el algoritmo Simplex y a entender mejor el vocabulario algebraico definido en este capítulo. Escojamos como punto de partida el punto extremo A. Como hemos mencionado anteriormente, el método Simplex se mueve de punto extremo a punto extremo adyacente siempre que el objetivo mejore. El punto A tiene dos puntos extremos adyacentes. Ambos mejoran el objetivo. Escogemos el que aumenta más el valor del objetivo (punto B). En el punto A teníamos una solución básica factible (dos variables con valor 0, X_1 y X_2, y las otras con valores positivos). Cuando pasamos al punto extremo B, observamos que una variable estrictamente positiva en A pasa a tener el valor 0 (la variable X_6), mientras que una de las variables con valor 0 pasa a tener un valor estrictamente positivo (la variable X_2). Este proceso se repite cada vez que pasamos de punto extremo a punto extremo adyacente: una de las variables básicas (con valor positivo) pasa a ser no-básica (valor 0) y una variable no-básica pasa a ser positiva (variable básica). En el punto D, dos variables que tenían valor positivo en el punto extremo adyacente anterior pasan a tener un valor 0. En otras palabras, dos soluciones básicas son adyacentes si todas, menos una de sus variables no-básicas, son las mismas. Entonces, pasar de una solución básica factible a una adyacente implica el cambio del estado básico de una variable a uno no básico, y viceversa.

En términos generales, el número de variables no básicas de una solución básica siempre es igual a los grados de libertad del sistema de ecuaciones de la forma canónica. El número de variables básicas siempre es igual al número de restricciones funcionales.

Propiedades de las soluciones factibles en un punto extremo

1. Si existe una única solución óptima, entonces ésta tiene que ser obligatoriamente una solución factible en un punto extremo (una solución básica factible). Si hay varias soluciones óptimas, entonces, como mínimo, tiene que haber dos que sean factibles en puntos extremos adyacentes.
2. Existe un número finito de soluciones factibles en los puntos extremos.
3. Si una solución en un punto extremo es igual o mejor (según el valor de Z) que todas las soluciones de los puntos extremos adyacentes, entonces ésta es igual o mejor que todas las otras soluciones en todos los puntos extremos; es decir, es óptima.

Ahora que ya conocemos los pasos que efectúa el método Simplex para buscar una solución óptima de un programa lineal, hace falta estudiar *cómo* se realizan éstos. Para entender la mecánica del método, tenemos que dar respuestas a las preguntas siguientes:

1. **Paso inicial:** ¿Cómo seleccionamos la solución factible inicial en un punto extremo (la solución básica factible inicial)?

2. **Paso iterativo:** Cuando buscamos un traslado a una solución factible en un punto extremo adyacente (una solución básica factible adyacente):

   a) ¿Cómo se selecciona la dirección del traslado? (¿Qué variable no básica se escoge para transformarla en básica?)
   b) ¿Adónde se realiza el traslado? (¿Qué variable básica se transforma en no-básica?)
   c) ¿Cómo identificamos la nueva solución?

3. **Prueba de optimalidad:** ¿Cómo determinamos que la solución factible en un punto extremo (solución básica factible) no tiene soluciones factibles en un punto extremo adyacente (soluciones básicas adyacentes) que sean mejores?

Para responder a estas preguntas, de momento consideraremos únicamente el caso de un programa lineal con restricciones de tipo «menor o igual» (≤). Más adelante ampliaremos el análisis cuando el problema también contiene los otros tipos de restricciones.

En primer lugar re-escribimos nuestro ejemplo en la forma canónica equivalente:

$$\text{Max } Z$$

$$\begin{array}{llllllll}
& \text{s. a} & & & & & & \\
(0) & Z - X_1 & -1{,}4X_2 & & & & & = 0 \\
(1) & X_1 & +0{,}5X_2 & +X_3 & & & & = 6 \\
(2) & 0{,}5X_1 & +X_2 & & +X_4 & & & = 6 \\
(3) & X_1 & +X_2 & & & +X_5 & & = 7 \\
(4) & 1{,}4X_1 & +X_2 & & & & +X_6 & = 9
\end{array}$$

$$X_1,\ X_2,\ X_3,\ X_4,\ X_5,\ X_6 \geq 0$$

Obsérvese que ahora la ecuación (0) del objetivo está incluida dentro del sistema de ecuaciones y que podemos considerar *Z* como una variable adicional.

1. *Paso inicial*

Este paso inicial consiste en encontrar cualquier solución básica factible. Una manera fácil de hacerlo es igualando las variables estructurales del modelo a 0. Si observamos la forma canónica equivalente tendremos que igualando X_1 y X_2 a 0 las variables de holgura automáticamente cogen valores positivos ($X_3 = 6$, $X_4 = 6$, $X_5 = 7$, $X_6 = 9$) correspondiente al punto extremo A. Por lo tanto, la solución factible será (0,0,6,6,7,9).

La razón por la cual la solución encontrada se deduce rápidamente es debido a que cada ecuación tiene una única variable básica con un coeficiente asociado a ella igual a + 1, y que esta variable básica no aparece en ninguna otra ecuación del sistema. Pronto observaremos que, cuando el conjunto de variables básicas cambia, el algoritmo Simplex utiliza un método algebraico llamado *eliminación de Gauss* para poner las ecuaciones en esta forma tan conveniente para obtener las soluciones básicas factibles subsecuentes. Esta forma funcional (una variable básica por ecuación con coeficiente + 1) se denomina *forma apropiada de eliminación gausiana.*

2) *Paso iterativo:*

En cada iteración, el método Simplex se mueve desde una solución básica factible a una solución básica factible adyacente que mejora el objetivo. Este movimiento consiste en convertir una variable no-básica (llamada *variable básica entrante*) en una variable básica y, al mismo tiempo, convertir una variable básica (llamada *variable básica saliente*) en variable no-básica, y a identificar la nueva solución básica factible.

Pregunta a): ¿Cuál es el criterio para seleccionar la variable básica entrante?
Las candidatas para la variable básica entrante son las n variables no básicas actuales. Esta variable, que escogeremos para pasar de no-básica a básica, pasará de tener un valor 0 a tener un valor positivo, mientras que las restantes seguirán con valor 0. Como el método Simplex requiere que este cambio implique una mejora en el objetivo, es necesario que la tasa de cambio en Z, al aumentar el valor de la variable básica entrante, sea positivo. Observemos la ecuación (0) del sistema. Esta expresión refleja el valor de Z en función de las variables no-básicas, y por lo tanto el coeficiente asociado a estas variables es la tasa de cambio del valor del objetivo. Si, por ejemplo, X_2 pasa de ser 0 a ser 1, el objetivo aumentará en 1,4 unidades. Como criterio, escogeremos la variable cuyo coeficiente aumente más el objetivo al pasar a ser básica [14].

En nuestro ejemplo, las dos variables no-básicas son candidatas a entrar en la base, ya que aumentarían el valor del objetivo. Escogemos X_2, ya que tiene un coeficiente en la ecuación (0) más elevado.

Pregunta b): ¿Cómo identificamos la variable básica saliente?
Si ignoramos las variables de holgura, al aumentar el valor de X_2 manteniendo X_1 igual a 0, nos desplazamos por el eje de las ordenadas (que corresponden precisamente a los valores de X_2). La solución adyacente se alcanza en el punto B, que viene determinada por la restricción $0{,}5\,X_1 + X_2 \leq 6$, que se cumplirá con igualdad y por lo tanto acotará en 6 el valor de X_2.

Cuando escribimos el problema en forma canónica, las soluciones factibles tienen que cumplir tanto las restricciones funcionales

---

[14] Este criterio es subjetivo y no implica que la solución óptima sea alcanzada más rápidamente.

como las de no-negatividad de todas las variables, incluidas las de holgura. Cuando vamos aumentando el valor de X_2 manteniendo $X_1 = 0$ (variable no-básica), algunas de las variables en la base actual (X_3, X_4, X_5, X_6) también van cambiando de valor para mantener válido el sistema de ecuaciones. Algunas de estas variables se reducirán al aumentar X_2. La solución básica adyacente se alcanza cuando la primera variable básica que tenía valor positivo pasa a ser igual a cero (recordemos las restricciones de no-negatividad). Esta variable será la que sale de la base y por lo tanto se transformará en no-básica. Por lo tanto, una vez escogida la variable que entrará en la base, la variable que sale de la base será aquella que llegue primero a 0. La variable básica actual con la cota superior más pequeña junto con la restricción su no-negatividad será la escogida.

Examinemos esta cuestión en nuestro ejemplo. Tenemos que las variables básicas candidatas a salir de la base son X_3, X_4, X_5 y X_6. En el Cuadro 3.2, se presentan los cálculos para identificar cuál es la variable básica saliente.

Como X_1 es una variable no-básica, tendremos que $X_1 = 0$ en la segunda columna del Cuadro 3.2. La tercera columna indica las cotas superiores para X_2 antes de que la variable básica correspondiente a la primera columna sea negativa. Por ejemplo, $X_3 = 0$ si $X_2 = 12$ (mientras que $X_3 > 0$ si $X_2 < 12$, y $X_3 < 0$ cuando $X_2 > 12$). Como X_4 (la variable de holgura correspondiente a la restricción $0{,}5\ X_1 + X_2 \leq 6$) impone la cota negativa más pequeña sobre X_2, la variable básica saliente será X_4, de manera que en la nueva situación tendremos que $X_4 = 0$ (no-básica) y $X_2 = 6$ (básica).

CUADRO 3.2
Cálculos para obtener la variable saliente

Variable básica	Ecuación	Cota Superior para X_2
X_3	$X_3 = 6 - X_1 - 0{,}5X_2$	$X_2 \leq 6/0{,}5=12$
X_4	$X_4 = 6 - 0{,}5X_1 - X_2$	$X_2 \leq 6 \leftarrow$ mínimo
X_5	$X_5 = 7 - X_1 - X_2$	$X_2 \leq 7$
X_6	$X_6 = 9 - 1{,}4X_1 - X_2$	$X_2 \leq 9$

Pregunta c): ¿Cómo podemos identificar de manera convincente la nueva solución básica factible?

Después de haber identificado las variables entrantes y salientes de la base (incluyendo el valor de la variable básica entrante), necesitamos conocer cuál es el valor nuevo del resto de variables básicas. Para poder calcular estos valores, el método Simplex utiliza la forma apropiada de eliminación de Gauss que teníamos en el paso inicial (aquella en la cual cada ecuación tiene únicamente una variable básica con coeficiente +1, y esta variable básica aparece en una única ecuación). Se trata de encontrar la nueva forma apropiada después del cambio de base. Se necesita realizar dos operaciones algebraicas normalmente utilizadas para resolver sistemas de ecuaciones lineales. Estas operaciones son:

1. Multiplicar (o dividir) una ecuación por una constante diferente de 0.
2. Sumar (o restar) un múltiple de una ecuación con otra ecuación.

Estas operaciones son legítimas porque implican únicamente: 1) multiplicar cosas iguales (los dos lados de la ecuación) por una constante y 2) sumar cosas iguales con cosas iguales. Por lo tanto, una solución que cumple un sistema de ecuaciones determinado también lo hará después de la transformación.

Vamos a ver cómo funciona en nuestro ejemplo. Consideremos un sistema de ecuaciones originales, en el cual se muestran las variables básicas en negrita. El problema se puede escribir de la forma siguiente:

$$
\begin{array}{lcrrrrrrl}
(0) & Z & -X_1 & -1{,}4X_2 & & & & & =0 \\
(1) & & X_1 & +0{,}5X_2 & +X_3 & & & & =6 \\
(2) & & 0{,}5X_1 & +X_2 & & +X_4 & & & =6 \\
(3) & & X_1 & +X_2 & & & +X_5 & & =7 \\
(4) & & 1{,}4X_1 & +X_2 & & & & +X_6 & =9
\end{array}
$$

Obsérvese que X_2 ha sustituido a X_4 como variable básica en la ecuación (2). Ahora tenemos que resolver este sistema de ecuaciones para encontrar los valores de las variable básicas X_2, X_3, X_5, X_6 (recordemos que ahora X_1 y $X_4 = 0$) y de Z. Como que X_2 tiene un coeficiente igual a +1 en la ecuación (2), no necesitamos realizar ninguna transformación. Si el coeficiente hubiese sido diferente de 1, hubiéramos tenido que dividir ambos lados de la ecuación por éste para que X_2 tenga un coeficiente asociado igual a 1. Nos referiremos al resultado de esta transformación Ecuación (2) nueva.

El paso siguiente es eliminar X_2 de las otras ecuaciones. Comencemos por la ecuación (0). Tenemos que realizar la operación siguiente:

$$\text{Ec. (0) nueva} = \text{ec. (0) antigua} + 1.4 \cdot \text{ec.(2) nueva}$$

Es decir:

$$
\begin{array}{rrrrl}
Z & -X_1 & -1{,}4X_2 & & =0 \\
+(& 0{,}7X_1 & +1{,}4X_2 & +1{,}4X_4 & =8{,}4) \\
\hline
 & Z\ -0{,}3X_1 & & +1{,}4X_4 & =8{,}4
\end{array}
$$

Tenemos que realizar el mismo procedimiento para las ecuaciones (1), (3) y (4). Lo haremos a continuación para la ecuación (1). Tenemos que, para eliminar X_2 de la ecuación (1), tenemos que realizar la operación siguiente:

$$\text{Ec.(0) nueva} = \text{ec. (0) antigua} - 0{,}5 \cdot \text{ec.(2) nueva}$$

$$
\begin{array}{rrrrl}
 & X_1 & +0{,}5X_2 & +X_3 & =6 \\
-(& 0{,}25X_1 & +0{,}5X_2 & +0{,}5X_4 & =3) \\
\hline
 & 0{,}75X_1 & & +X_3\ -0{,}5X_4 & =3
\end{array}
$$

Si utilizamos el mismo procedimiento para las ecuaciones (3) y (4) para eliminar X_2 y obtenemos la nueva forma gausiana siguiente del sistema de ecuaciones:

$$
\begin{array}{llllllll}
(0) & Z\ -0{,}3X_1 & & & +1{,}4X_4 & & & =8{,}4 \\
(1) & \quad 0{,}75X_1 & & +X_3 & -0{,}5X_4 & & & =3 \\
(2) & \quad 0{,}5X_1 & +X_2 & & +X_4 & & & =6 \\
(3) & \quad 0{,}5X_1 & & & -X_4 & +X_5 & & =1 \\
(4) & \quad 0{,}9X_1 & & & -X_4 & & +X_6 & =3
\end{array}
$$

Si comparamos este nuevo sistema de ecuaciones con el anterior, veremos que sigue teniendo la forma apropiada de eliminación de Gauss que permite obtener inmediatamente el valor de las variables en la solución (recordemos que $X_4 = 0$ y $X_1 = 0$). Hay que observar que en la ecuación (0) siempre están únicamente las variables no-básicas. Ahora tenemos una nueva solución básica factible igual a (0,6,3,0,1,3) que corresponde al punto extremo B. El valor del objetivo es igual a 8,4.

El siguiente paso es ver si esta nueva solución es la óptima. Para ello examinamos en la siguiente ecuación [que corresponde a la ecuación (0)] los coeficientes de las variables no básicas:

$$Z = 8{,}4 + 0{,}3X_1 - 1{,}4X_4$$

Como la no-básica X_1 tiene un coeficiente positivo (0,3), si la variable pasa a ser positiva el valor del objetivo aumentará. Por lo tanto, no estamos en la solución óptima y hay que realizar de nuevo el proceso, en donde X_1 entrará en la base y otra variable básica dejará de serlo.

Segunda iteración

- **Paso 1.** Como que la ecuación (0) actual es $Z = 8{,}4 + 0{,}3X_1 - 1{,}4X_4$, la función sólo aumentará si X_1 aumenta. Ya tenemos la variable que entrará en la base.

- **Paso 2.** El límite superior sobre X_1 antes de que las variables básicas sean negativas está indicado en el Cuadro 3.3:

CUADRO 3.3
Cálculos para obtener la variable saliente

Variable básica	Ecuación	Cota Superior para X_2
X_3	$X_3 = 3 - 0{,}75X_1 - 0{,}5\,X_4$	$X_1 \leq 3/(0{,}75) = 4$
X_2	$X_2 = 6 - 0{,}5X_1 - X_4$	$X_1 \leq 12$
X_5	$X_5 = 1 - 0{,}5X_1 + X_4$	$X_1 \leq 2$ ← mínimo
X_6	$X_6 = 3 - 0{,}9X_1 + X_4$	$X_1 \leq 3{,}3$

Escogeremos X_5 como la variable básica saliente, ya que es la que, a medida que aumenta el valor de X_1, alcanza primero el valor 0.

- **Paso 3.** Después de eliminar X_1 de todas las ecuaciones del conjunto actual, excepto de la ecuación (3), en la cual X_1 sustituye a X_5 como variable básica, el nuevo conjunto de ecuaciones en la forma apropiada de eliminación de Gauss es:

$$
\begin{array}{llllllll}
(0) & Z & & & +0{,}8X_4 & +0{,}6X_5 & & =9 \\
(1) & & & +X_3 & +X_4 & -1{,}5X_5 & & =1{,}5 \\
(2) & & +X_2 & & +2X_4 & -X_5 & & =5 \\
(3) & X_1 & & & -2X_4 & +2X_5 & & =2 \\
(4) & & & & -0{,}8X_4 & +1{,}8X_5 & +X_6 & =1{,}2
\end{array}
$$

La solución básica factible siguiente es (2; 5; 1,5; 0; 0; 5) y el valor del objetivo es $Z = 9$.

- **Prueba de optimalidad.**

Tenemos que verificar si las variables no-básicas del objetivo tienen coeficientes que permitan aumentar el valor del objetivo si éstas cogen valores positivos. El nuevo objetivo es:

$$Z = 9 - 0{,}8X_4 - 0{,}6X_5$$

Como las variables no-básicas tienen un coeficiente negativo, el valor del objetivo no puede aumentar. Estamos en el óptimo y la solución final es $X_1 = 2$, $X_2 = 6$ y $Z = 9$.

En resumen, el método Simplex tiene los pasos siguientes:

1. Introducir las variables de holgura para obtener la forma canónica del programa.

2. Encontrar una solución inicial de un punto extremo y realizar la prueba de optimalidad.

3. Si no estamos en el óptimo:

   a) Determinar la variable básica entrante: seleccionar la variable no básica que, al aumentar su valor, aumente más rápidamente el valor del objetivo.

   b) Determinar la variable básica saliente: ésta es la que alcanza el valor 0 más rápidamente a medida que aumentamos la variable entrante.

   c) Una vez que sabemos cuál es la variable básica que sale de la base, se determina la nueva solución básica factible: a partir del conjunto actual de ecuaciones se aíslan las variables básicas y Z en términos de las variables no-básicas utilizando el método de eliminación de Gauss. Las variables no-básicas se igualan a 0; cada variable básica junto con Z es igual al nuevo lado derecho de la ecuación en la cual aparece con coeficiente +1.

4. Examinamos si la nueva solución encontrada es óptima: únicamente necesitamos examinar los coeficientes de las variables no básicas que están en el objetivo. Si todos los coeficientes son negativos, estamos en el óptimo. Por otro lado, si como mínimo uno de los coeficientes asociados a las variables básicas es positivo, tenemos que repetir los pasos 2 y 3.

3.3. Adaptación a otro tipo de modelos

Hasta ahora hemos estudiado el método Simplex para problemas de maximización con restricciones con la desigualdad $\leq$. Pero hay otros casos como los problemas de minimización y la existencia de restricciones con igualdad o con desigualdad $\geq$. A continuación veremos cómo adaptar la formulación del modelo con alguna de estas características para poder utilizar el método Simplex.

3.3.1. *Restricciones con igualdad*

El problema básico con las restricciones de igualdad es la obtención de una solución básica factible inicial. Supongamos que, en nuestro ejemplo, la restricción (3) se tiene que cumplir con igualdad ($X_1 + X_2 = 7$). En este caso, en principio, no hay que introducir una variable de holgura para formar la forma canónica.

Si procedemos a encontrar una solución inicial factible igualando X_1 y X_2 a 0 nos encontramos con el problema de que la restricción (3) no se cumple. Para poder obtener una solución inicial factible, nos vemos obligados a introducir una nueva variable no-negativa, denominada *artificial* S_3, de la siguiente forma:

$$X_1 + X_2 + S_3 = 7$$

Gracias a la introducción de esta variable artificial S_3 ya podemos encontrar una solución inicial factible en donde S_3 es una variable básica igual a 7. De hecho, hemos aumentado el número de variables añadiendo una que no tiene ninguna interpretación económica, pero que nos sirve para encontrar una solución inicial factible. Es meramente un artificio matemático. Pero, en la solución final, queremos que esta solución tenga el valor 0 (sea no básica), ya que, si esto no es así, el problema no tendría sentido (la restricción no se cumpliría con igualdad). Para poder conseguirlo, añadimos esta variable artificial en el objetivo, pero con un coeficiente negativo de valor muy elevado (respecto a los otros), que llamaremos M:

$$Z = X_1 + 1{,}4X_2 - MS_3$$

Como este valor penaliza la variable en el objetivo, el método Simplex escogerá esta variable para salir de la base y nunca más volverá a entrar (es decir, se quedará con el valor 0). Por lo tanto, en la solución final S_3 tendrá el valor 0. Si esto no fuera así, el problema sería *infactible.* En la próxima sección veremos cómo eliminar esta variable del objetivo.

3.3.2. *Restricciones con dirección* ≥

Supongamos ahora que la restricción (3) es la siguiente:

$$X_1 + X_2 \geq 7$$

En este caso tenemos que encontrar una solución inicial de la misma forma que hacíamos anteriormente para poder ejecutar el método Simplex. Ahora bien, en este caso añadimos una *variable de exceso* no-negativa, E_3, que mide la diferencia entre el valor del lado izquierdo de la ecuación ($X_1 + X_2$) y el lado derecho (7). Esta variable tendrá un signo negativo en la ecuación:

$$X_1 + X_2 - E_3 = 7$$

Ahora bien, al fijar inicialmente X_1 y X_2 iguales a 0, E_3 se igualará a –7, por lo que tendrá un valor negativo, incumpliendo las condiciones de no-negatividad de todas las variables en el método Simplex. De nuevo, tenemos que recurrir al artificio de introducir una variable artificial que nos permita obtener una solución factible inicial S_3:

$$X_1 + X_2 - E_3 + S_3 = 7$$

En este caso escogemos S_3 como variable básica inicial correspondiente a la restricción (3). Como en el caso anterior (restricciones con igualdad), añadiremos la variable artificial S_3 en el objetivo con un coeficiente –M. Si esta variable continúa con valor positivo al final del método Simplex, el problema es infactible.

El hecho de añadir la variable artificial en el objetivo implica que, al iniciar el método Simplex, el cuadro inicial no esté en la forma apropiada de eliminación gausiana, ya que esta forma requiere que todas las variables básicas tengan un coeficiente 0 en la ecuación (0) correspondiente al objetivo, y en este caso la variable básica S_3 tiene un coeficiente igual a –M. Entonces, para poder iniciar el método Simplex, tanto si tenemos restricciones con igualdad o desigualdad ≥, tenemos que transformar esta ecuación (0) en la forma apropiada de eliminación de Gauss, para poder así determinar tanto la variable que entrará en la base como el test de optimalidad. De nuevo, el procedimiento es el de siempre: el método de eliminación de Gauss. En este caso, el procedimiento es muy similar al utilizado hasta ahora en el método Simplex. Tendremos que realizar la operación siguiente:

$$\text{Ec. (0) nueva} = \text{ec. (0) antigua} - \cdot \text{ec.(3)}$$

Es decir:

$$\begin{array}{rccccr} & Z & X_1 & -1.4X_2 & & +MS_3 = 0 \\ -M(& & X_1 & +X_2 & -E_3 & S_3 = 7) \\ \hline & Z & (-M+1)X_1 & (-M-1{,}4)X_2 & -ME_3 & = -7M \end{array}$$

Ahora ya podemos proceder con el método Simplex, ya que todas las variables básicas en la ecuación (0) tienen un coeficiente asociado igual a 0. Ahora tenemos que decidir qué variable no-básica tiene que entrar en la base. Escogeremos aquella cuyo coeficiente aumente más el objetivo. En este caso, escogeríamos E_3 como variable básica entrante y procederíamos a buscar la variable básica saliente de la misma forma que lo hicimos anteriormente. Hemos de observar que cuando E_3 entra en la base y otra variable sale de la base (es decir, nos desplazamos a un nuevo punto extremo adyacente), el coeficiente de E_3 en el objetivo tomará el valor 0. A medida que el procedimiento continúa, las variables con el valor M en el objetivo van entrando en la base y llegará un punto en que M desaparecerá del sistema. Si en la solución final aún tenemos M en la ecuación (0), el sistema no tiene solución.

Si tenemos más de una restricción con igualdad, el procedimiento es exactamente el mismo. Cada una de las variables artificiales tendrá un coeficiente –M en el objetivo y tendremos que encontrar la forma apropiada de eliminación de Gauss.

3.3.3. *Minimización*

Hasta ahora, hemos examinado el método Simplex cuando estamos maximizando el objetivo. Pero, en muchos casos, tenemos que minimizar el objetivo (por ejemplo, minimizar costes, minimizar el grado de contaminación o minimizar la mortalidad). Lo más sencillo es multiplicar el objetivo por –1. Por ejemplo:

$$\text{Min } Z = 3X_1 + 4X_2$$

Es equivalente a:

$$\text{Max } -Z = -3X_1 - 4X_2$$

Una vez hecha esta transformación, podemos aplicar el método Simplex descrito en esta sección. La causa de esta equivalencia es que, cuanto menor es Z, mayor es –Z.

Otra manera de operar con un objetivo de minimización es seleccionar la *variable no-básica entrante que reduzca* en mayor grado el valor del objetivo.

3.3.4. *Variables no acotadas*

Puede ocurrir que en algunas formulaciones las variables puedan coger valores negativos. En este caso, hay que modificar el modelo para poder utilizar el método Simplex, ya que éste únicamente permite que las variables tomen valores positivos o cero.

Supongamos que la variable X_i no está acotada inferiormente. Para poder resolver el problema, tendremos que sustituir esta variable en todas las ecuaciones por dos variables X_i^+ y X_i^- de la manera siguiente:

$$X_i = X_i^+ - X_i^-$$

en donde $X_i^+ \geq 0$ y $X_i^- \geq 0$. Como estas dos variables pueden coger cualquier valor no-negativo, su diferencia puede ser cualquier valor (positivo o negativo). Ahora ya podemos aplicar el método Simplex. En la solución final, debido a las propiedades geométricas de la solución factible en un punto extremo, nunca tendremos las dos variables con valores positivos. O únicamente una de ellas tiene valor estrictamente positivo y la otra igual a 0 (o viceversa), o las dos son iguales a 0.

3.4. Situaciones especiales en el método Simplex

¿Qué pasa cuando vamos a escoger la variable no-básica entrante y hay un empate en el criterio? ¿Cómo detectamos problemas sin solución? ¿y si la solución es infinita? A continuación examinaremos cómo el método Simplex lidia con estas situaciones.

Empate en la variable entrante
Si hay dos variables que tienen el coeficiente más grande (en valor absoluto) igual en la ecuación (0), se escoge arbitrariamente una de ellas para entrar en la base.

Empate en la variable saliente
Supongamos que ahora el empate se produce entre dos o más variables básicas al examinar el criterio de salida. Si esto sucede, to-

das las variables alcanzan el valor 0 al mismo tiempo cuando aumenta el valor de la variable entrante. Entonces, las variables básicas que no habíamos escogido como salientes de la base también tendrán valor 0 en la solución. Este tipo de soluciones se llaman *degeneradas*. Incluso, si una de estas variables continúa con el valor 0 hasta que se selecciona como variable saliente en una iteración posterior, la variable no-básica entrante también se quedará con valor 0 y el valor del objetivo no cambiará. Puede pasar que, si Z se queda igual, en vez de mejorar el objetivo en cada iteración, el método Simplex entre en un ciclo que repite periódicamente las mismas soluciones, en vez de ir cambiando para aumentar el valor del objetivo. De hecho, se han elaborado programas lineales con ciclos infinitos. Por suerte, en la práctica esta situación es casi inexistente y normalmente los empates se rompen arbitrariamente.

No hay variable básica saliente: Z no acotado
¿Qué pasa cuando no hay variables básicas candidatas a salir en la base? O, en otras palabras, ¿qué hacemos cuando todos los cocientes calculados para seleccionar la base son de tal manera que no hay ninguno en positivo? Recordemos que, a medida que aumentábamos el valor de la variable no-básica entrante, había como mínimo variable básica que iba disminuyendo hasta llegar a tener un valor 0, que determinaba automáticamente el nuevo valor de la variable entrante. Pues bien, puede haber situaciones en donde a medida que aumenta el valor de la variable entrante todas las variables básicas *también aumentan de valor* (o no cambian). Simplemente, el problema tiene una solución *infinita,* ya que no hay ninguna restricción que acote el objetivo.

Soluciones óptimas múltiples
Como hemos visto, el método Simplex se para cuando encuentra una solución óptima. Pero, como hemos visto en el método gráfico, puede haber situaciones en las que hay soluciones óptimas múltiples... Siempre que el problema tiene más de una solución óptima factible, como mínimo una variable no-básica tiene el coeficiente igual a 0 en la ecuación (0) final, de manera que si su valor aumenta, Z no cambia. Si esta situación aparece, podemos encontrar otra solución óptima introduciendo esta variable no-básica en la base. Así podemos encontrar otras soluciones que, sin cambiar el valor del objetivo, nos ayuden a tomar una decisión en función del valor de las variables en el óptimo.

3.5. Soluciones con ordenador

Por ahora hemos visto tres métodos para encontrar soluciones de programas lineales. Pero todos ellos son muy ineficientes si se tiene que hacer los cálculos con lápiz y papel incluso para problemas pequeños. Actualmente, existen un sinfín de programas de ordenador que resuelven problemas lineales muy eficientemente, incluso programas con miles de variables y restricciones. Los programas de hoja de cálculo también están incorporando métodos para obtener soluciones de programas lineales. En esta sección describiremos cómo programar y solucionar un modelo de programación lineal en la hoja de cálculo Excel 97® de Microsoft [15]. Utilizaremos mismo ejemplo de las secciones anteriores. También supondremos que se tienen conocimientos básicos de funcionamiento de este programa.

La formulación es:

$$\begin{aligned} \text{Max } Z = \; & X_1 + 1{,}4X_2 \\ \text{s. a.} \; & \\ & X_1 + 0{,}5X_2 \le 6 \\ & 0{,}5X_1 + X_2 \le 6 \\ & X_1 + X_2 \le 7 \\ & 1{,}4X_1 + X_2 \le 9 \\ & X_1, X_2 \ge 0 \end{aligned}$$

Obsérvese que el conjunto de restricciones puede representarse de formal matricial sólo con los coeficientes de las variables:

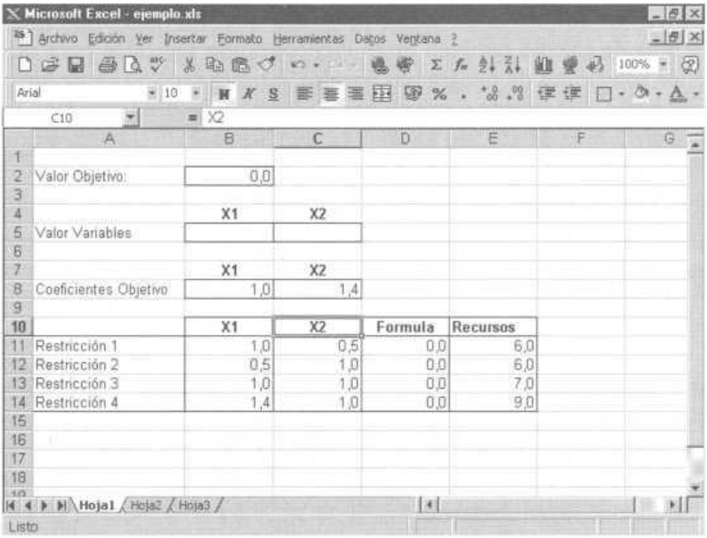
1	0,5	≤	6
0,5	1	≤	6
1	1	≤	7
1,4	1	≤	9

La primera columna corresponde a los coeficientes de X_1 y la segunda a los coeficientes de X_2. Precisamente vamos a escribir esta matriz en las celdas de la hoja de cálculo. En la Figura 3.5 hemos escrito el planteamiento del problema.

En los rangos B11-B14 y C11-C14 hemos escrito los coeficientes de X_1 y X_2 en las restricciones. En el rango E11-E14 figuran los valores de los recursos (lado derecho de las restricciones y en las

[15] La versión que se utiliza en este apartado corresponde a Office 97, aunque en versiones anteriores también existe el módulo de programación lineal.

FIGURA 3.5
Ejemplo en Excel



celdas B8 y C8 los coeficientes de las variables en el objetivo. Ahora tenemos que escribir las fórmulas correspondientes a las restricciones y a la función objetivo.

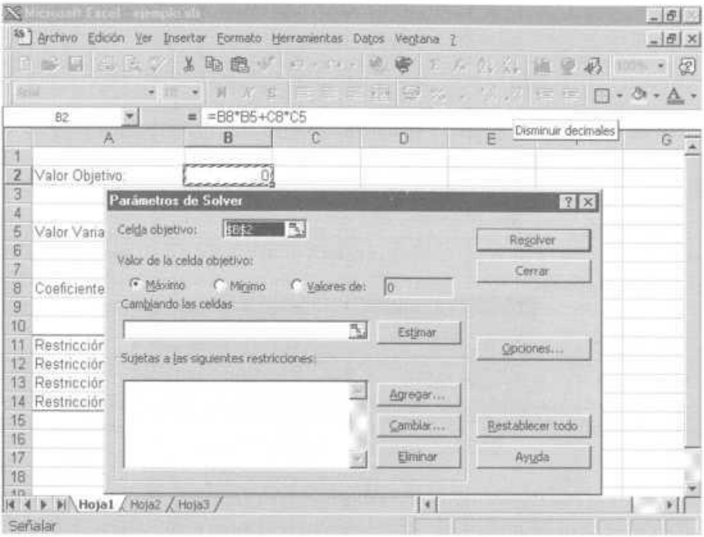
Las celdas B5 y C5 representarán los valores de las variables de decisión X_1 y X_2. La fórmula de la función objetivo está escrita en la celda B2. La fórmula es la siguiente: =B8*B5 + C8*C5. Las fórmulas del lado izquierdo de las restricciones están escritas en el rango D11-D14. Estas son:

=B11*B5 + C11*C5
=B12*B5 + C12*C5
=B13*B5 + C13*C5
=B14*B5 + C14*C5

Ahora ya tenemos preparado el modelo. Obsérvese que por el momento las celdas con fórmulas tienen el valor 0. Esto es debido a que por ahora las celdas asociadas a las variables de decisión están vacías.

El siguiente paso es indicar a la hoja de cálculo dónde está el problema. Entramos en la opción **Herramientas** y escogemos en el menú el **Solver.** Entonces aparecerá un recuadro como el de la Figura 3.6.

FIGURA 3.6
Cuadro de la opción Solver



Ahora tenemos que indicar las celdas en donde están las fórmulas. En la casilla «Celda objetivo» ponemos la referencia de la celda en donde está la función objetivo (B2). Luego indicamos que es un problema de maximización. Las referencias de las variables se indican en el recuadro «Cambiando las celdas» (B5;C5). Finalmente tenemos que introducir las restricciones. Para ello entramos en la opción Agregar y saldrá el recuadro de la Figura 3.7. En él tenemos que indicar dónde está el lado izquierdo (la fórmula) de cada restricción, el signo de la desigualdad y el lado derecho de cada restricción. Cada vez que entramos una restricción adicional escogemos la opción agregar. Después de haber entrado las restricciones correspondientes, Excel también exige poner las restricciones de no-negatividad. Para ello seguimos agregando dos restricciones, una para cada variable. Para ello tenemos que indicar en el lado izquierdo la referencia de la celda correspondiente a la variable de decisión, y en el lado derecho pondremos el valor 0, habiendo escogido previamente el sentido de la desigualdad (≥). La Figura 3.8 muestra el resultado final de introducir el problema.

Ahora ya podemos resolver el problema. Escogemos la opción **Resolver** y al cabo de unos breves momentos saldrá una pantalla indicando que la solución ha sido encontrada. La solución óptima de las variables de decisión y el valor del objetivo ahora aparecen en las celdas (ver Figura 3.9). La opción Solver también permite obtener automáticamente informes sobre la solución final.

FIGURA 3.7
Introducción de las restricciones

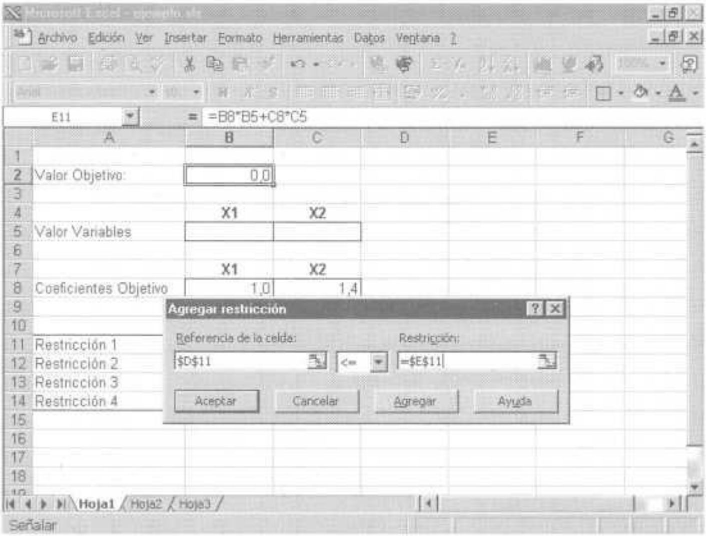


FIGURA 3.8
Resultado final de la programación

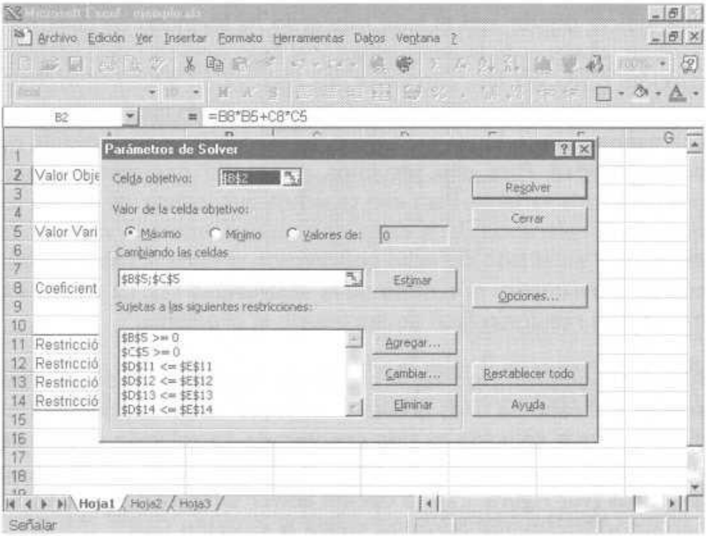
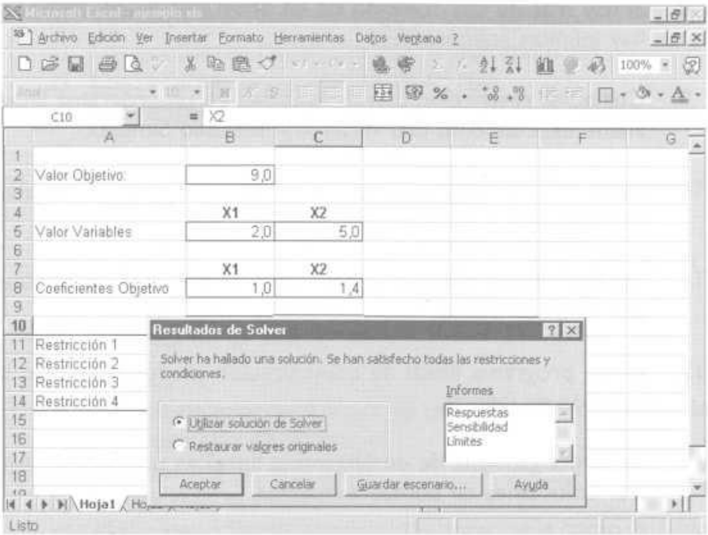


FIGURA 3.9
Solución óptima del ejemplo



3.6. Ejercicios

Cierto/falso

1. En un programa lineal, todas las funciones que forman el conjunto de restricciones y el objetivo son lineales.
2. El método gráfico es muy útil porque permite encontrar soluciones de cualquier programa lineal.
3. Cualquier solución que cumple, como mínimo, una restricción de un programa lineal pertenece a la región factible.
4. Una solución óptima no utiliza necesariamente todos los recursos disponibles en el problema.
5. La intersección de dos restricciones cualquiera es un punto extremo de la región factible.
6. Una restricción con igualdad en general acota más la región factible que una con desigualdad.

7. Una solución de un programa lineal siempre se encuentra en un punto extremo.
8. Si hay más de una solución de un programa lineal, entonces hay infinitas soluciones.
9. Cualquier restricción con desigualdad en un programa lineal añade exactamente una variable al método Simplex.
10. Cada solución factible encontrada por el algoritmo Simplex corresponde a un punto extremo.
11. Las variables artificiales se añaden al modelo para encontrar una solución inicial.
12. En un programa lineal de maximización sin restricciones, la solución es infinita.
13. La forma gausiana de un programa lineal es útil para encontrar soluciones en puntos extremos.
14. Si hay un empate en el criterio de escoger la variable básica saliente, la solución es infinita.
15. Si no tenemos ninguna variable candidata a salir de la base, la solución es no acotada.
16. Si una variable básica tiene el coeficiente del objetivo igual a 0 en una iteración del método Simplex antes de llegar al óptimo, entonces hay infinitas soluciones óptimas.
17. En una restricción con igualdad añadimos una variable de holgura.
18. Para identificar una variable básica saliente, hay que saber antes qué variable no-básica va a entrar en la base.
19. Una variable no-básica es siempre igual a 0 en cualquier solución.
20. En la solución final de un programa lineal todas la variables básicas siempre tienen valores no-negativos.

Elección múltiple

1. ¿Qué no es esencial en un programa lineal?
   a) Tenemos que tener un objetivo bien definido.
   b) Los problemas tienen que ser de maximización.
   c) Los recursos tienen que ser limitados.
   d) Las variables son continuas.

2. En la programación lineal, no-negatividad implica que una variable no puede tener:
   a) Un coeficiente negativo en la función objetivo.
   b) Un coeficiente negativo en las restricciones.
   c) Un valor fraccional.
   d) Ninguna de las anteriores.

3. Si tenemos un programa lineal de maximización con todas las restricciones con dirección $\geq$:
   a) El problema no es factible.
   b) La solución es infinita.
   c) El algoritmo Simplex hace tantas iteraciones como variables en el problema.
   d) Ninguna de las anteriores.

4. La intersección de las restricciones de un programa lineal forma:
   a) Un conjunto convexo.
   b) El espacio factible de soluciones.
   c) Los puntos extremos.
   d) Todas las anteriores.

5. Para poder encontrar una solución inicial de un programa lineal, las restricciones con dirección $\geq$ se han de transformar en igualdad añadiendo:
   a) Una variable de exceso.
   b) Una variable artificial.
   c) Una variable de holgura y una artificial.
   d) Ninguna de las anteriores.

6. Supongamos un programa lineal con dos variables y una única restricción. Si estamos maximizando y la restricción es $\leq$:
   a) Las dos variables tendrán valores positivos.
   b) En general, una variable será positiva y la otra cero.
   c) El problema no tiene solución.
   d) Ninguna de las anteriores.

7. Supongamos un programa lineal con dos restricciones ≤, una restricción con = y tres variables. Cuando construimos la forma canónica para encontrar una solución inicial, tendremos un total de:
   a) 3 variables estructurales, 2 de holgura y 1 de exceso.
   b) 3 variables estructurales, 2 de exceso y 1 artificial.
   c) 3 variables estructurales, 2 de holgura y 1 artificial.
   d) Ninguna de las anteriores.

8. Si tenemos la solución óptima de un programa lineal con dos variables de decisión, ¿cuál de las opciones es la correcta?
   a) El problema tiene una única solución.
   b) La solución óptima se encuentra en un punto extremo o a lo largo de una recta que conecta dos puntos extremos.
   c) Todos los recursos se han consumido en la solución óptima.
   d) Todas las anteriores.

9. En una solución básica factible:
   a) Todas las variables son positivas.
   b) Estamos en el óptimo.
   c) Las variables no-básicas son 0.
   d) Ninguna de las anteriores.

10. En un punto extremo de la región factible:
    a) Hay una solución básica factible.
    b) Coinciden dos o más restricciones.
    c) Podemos encontrar el óptimo.
    d) Todas las anteriores.

11. En cada iteración del algoritmo Simplex, tenemos que:
    a) Como mínimo, una de las variables que tienen un valor positivo en la solución anterior coge el valor 0.
    b) Una variable no-básica se transforma en básica.
    c) El valor del objetivo mejora.
    d) Todas las anteriores.

12. Si en la solución óptima la variable de holgura X, correspondiente a una restricción con dirección ≤ y lado derecho (recurso) R, está en la base, entonces:
    a) El problema no es factible.
    b) No se ha consumido todo el recurso R.
    c) Se ha consumido todo el recurso R.
    d) Se puede obtener un solución mejor si el valor de R aumenta.

13. Si tenemos una solución óptima degenerada:
    a) Hay soluciones alternativas óptimas.

b) La solución no sirve de nada.
c) La solución no es factible.
d) Ninguna de las anteriores.

14. Si tenemos que en la ecuación (0) de la solución final del método Simplex hay el coeficiente –M asociado a una variable artificial, sabemos que el problema es:
a) Un problema de maximización.
b) Un problema de minimización.
c) No es factible.
d) Ninguna de las anteriores.

15. Si hay un valor negativo en los coeficientes de la ecuación (0) del método Simplex, sabemos que:
a) La solución es óptima.
b) Hemos cometido un error.
c) El problema no tiene límites.
d) Ninguna de las anteriores.

Problemas

3.1. Considerar el programa lineal siguiente:

$$\text{Max } Z = X_1 + 2X_2$$

$$\text{s. a.}$$

$$\begin{aligned} 2X_1 + 8X_2 &\leq 16 \\ X_1 + X_2 &\leq 5 \\ X_1, X_2 &\geq 0 \end{aligned}$$

a) Utilizar un método gráfico para encontrar una solución
b) Cambiar la función objetivo por $Z = X_1 + 6X_2$ y volver a solucionar el problema.
c) ¿Cuántos puntos extremos tiene la solución factible? Encontrar los valores de X_1 y X_2 en cada punto extremo.

3.2. Supongamos el programa lineal siguiente:

$$\text{Max } Z = 4X_1 + X_2 + X_3$$

$$\text{s. a.}$$

$$\begin{aligned} X_1 + 4X_2 + X_3 &\leq 12 \\ 5X_1 - 2X_2 + 4X_3 &\leq 11 \\ 2X_1 + 3X_2 + 3X_3 &\leq 20 \\ X_1, X_2, X_3 &\geq 0 \end{aligned}$$

Encontrar dos puntos extremos factibles.

3.3. Utilizar el método Simplex algebraico para solucionar el problema siguiente:

$$\text{Max } Z = -X_1 + X_2 + 2X_3$$

s. a.

$$\begin{aligned} 2X_1 + 8X_2 - X_3 &\leq 20 \\ -2X_1 + 4X_2 + 2X_3 &\leq 60 \\ 2X_1 + 3X_2 + X_3 &\leq 50 \\ X_1, X_2, X_3 &\geq 0 \end{aligned}$$

3.4. Considerar el problema siguiente:

$$\text{Max } Z = 2X_1 + 4X_2 + 3X_3$$

s. a.

$$\begin{aligned} X_1 + 3X_2 + 2X_3 &\leq 30 \\ X_1 - X_2 + X_3 &\leq 24 \\ 3X_1 + 5X_2 + 3X_3 &\leq 60 \\ X_1, X_2, X_3 &\geq 0 \end{aligned}$$

Sabemos que, en la solución óptima, X_1, X_2, $X_3 > 0$.

a) Describir cómo se puede utilizar esta información para adaptar el método Simplex de manera que el número de iteraciones sea mínimo (comenzando en la solución básica factible inicial normal). No vale hacer ninguna iteración.

b) Utilizar el procedimiento desarrollado para solucionar el problema.

3.5. Considerar el problema siguiente:

$$\text{Max } Z = X_1 + X_2 + X_3 + X_4$$

s. a.

$$\begin{aligned} X_1 + X_2 &\leq 3 \\ X_3 + X_4 &\leq 2 \\ X_1, X_2, X_3 &\geq 0 \end{aligned}$$

Utilizar el método Simplex para encontrar todas las soluciones factibles óptimas.

3.6. Considerar el programa siguiente:

$$\begin{array}{llllll} \text{Max } Z = & 2X_1 & -4X_2 & +5X_3 & -6X_4 & \\ \text{s. a.} & & & & & \\ & X_1 & +4X_2 & -2X_3 & +8X_4 & \leq 3 \\ & -X_1 & +2X_2 & +3X_3 & +4X_4 & \leq 2 \\ & \multicolumn{5}{l}{X_1, X_2, X_3, X_4 \geq 0} \end{array}$$

Determinar el número máximo de soluciones básicas posibles y la solución básica factible óptima.

3.7. Formular y encontrar la solución óptima del problema 2.1.

3.8. Formular y encontrar la solución óptima del problema 2.2.

3.9. Formular y encontrar la solución óptima del problema 2.3.

3.10. Formular y encontrar la solución óptima del problema 2.4.

3.11. Formular y encontrar la solución óptima del problema 2.5.

4. PROGRAMACIÓN LINEAL ENTERA

4.1. Introducción

Los modelos de programación lineal consideran que las variables de decisión son continuas, es decir, que pueden tomar en la solución final valores fraccionados. Pero, en muchos casos, una solución óptima de un programa lineal puede ser inservible si presenta fracciones. Supongamos, por ejemplo, que hemos construido un modelo para asignar personal médico a departamentos dentro de un hospital. En este caso, las variables de decisión (asignar personas a departamentos) tienen que ser enteras en la solución final. ¡No tendría sentido una solución en la cual 2,3 médicos fuesen asignados a la sección de dermatología!

Para poder encontrar soluciones de problemas en los cuales algunas o todas las variables tienen que ser enteras, se utiliza la *programación entera,* que no es más que una extensión de la programación lineal.

Otro tipo de modelos entran dentro de la programación entera binaria, que es un caso especial en donde todas o algunas de las variables representan acciones binarias, es decir, «hacer o no hacer». En este caso, las variables únicamente pueden adoptar los valores 0 ó 1. Este tipo de problemas es muy común en la toma de decisiones, en donde muchas veces tenemos que decidir si, por ejemplo, tenemos que construir un nuevo centro, si tenemos que invertir en un nuevo departamento, o si tenemos que modificar una estrategia de planificación de un servicio.

Cuando nos encontramos con este tipo de problemas, la formulación matemática no se ve alterada; únicamente en las restricciones de no-negatividad hay que indicar qué variables tienen que tomar valores enteros. El problema reside en encontrar soluciones que sean factibles, ya que el algoritmo Simplex no garantiza una solución adecuada al problema. En este capítulo examinaremos en primer lugar cómo podemos modificar el algoritmo Simplex para poder obtener soluciones óptimas. A continuación examinaremos algunos problemas de programación entera cuyas variables de decisión son binarias (decisiones «hacer o no hacer»).

4.2. El algoritmo de bifurcación y acotamiento

El Algoritmo de Bifurcación y Acotamiento [16] (ABA) se basa en el algoritmo Simplex para poder obtener soluciones enteras. En primer lugar, se aplica el algoritmo Simplex para obtener una solución inicial. Si en la solución obtenida al final del algoritmo Simplex todas las variables especificadas como enteras tienen valores enteros, no hace falta seguir, ya que se ha obtenido el óptimo; en caso contrario, es necesario aplicar el ABA. Básicamente, en cada iteración del ABA se escoge una variable que presenta una solución no-entera y se divide el problema en dos sub-problemas, añadiendo en cada uno de ellos una nueva restricción que acota esta variable por su valor entero superior en un caso, y por el valor su valor inferior entero por el otro. Cada sub-problema se resuelve con el método Simplex y se verifica si la solución es entera. En caso contrario, se vuelve a bifurcar el sub-problema en otros dos y se sigue procediendo hasta que se encuentra una solución entera. El proceso se realiza en todas las ramificaciones del árbol. Aunque este algoritmo pueda parecer complejo, el proceso es bastante sencillo. A continuación examinaremos con un ejemplo el ABA.

Supongamos que tenemos que encontrar la solución al problema lineal siguiente:

$$
\begin{aligned}
\text{Max } Z = \quad & X_1 + 1{,}4X_2 \\
\text{s.a.} \quad & \\
& X_1 + 0{,}5X_2 \le 6 \\
& 0{,}5X_1 + X_2 \le 5{,}5 \\
& X_1 + X_2 \le 6{,}8 \\
& 1{,}4X_1 + X_2 \le 9 \\
& X_1, X_2 \text{ enteras}
\end{aligned}
$$

En primer lugar utilizamos el método Simplex para obtener una solución del programa lineal relajado (sin considerar las restricciones

[16] En inglés, «branch and bound algorithm».

que fijan las variables como enteras). La solución obtenida es $Z = 8{,}5$; $X_1 = 2{,}6$ y $X_2 = 4{,}2$. Tenemos que las dos variables ofrecen soluciones fraccionadas. Hay que aplicar el ABA.

Definamos el problema original como P0. Escogemos X_1 y creamos dos sub-problemas P1 y P2 a partir del programa original. El primer sub-problema, P1, consistirá en el programa original P0 más la restricción $X_1 \leq 2$. El segundo sub-problema, P2, consistirá en el programa original más la restricción $X_1 \geq 3$. Es decir, estamos diciendo que X_1 no puede coger valores entre dos y tres. Solucionamos P1 y P2.

P1 = P0 + $X_1 \leq 2$. Solución: Z1 = 8,3; $X_1 = 2$ y $X_2 = 4{,}5$
P2 = P0 + $X_1 \geq 3$. Solución: Z2 = 8,3; $X_1 = 3$ y $X_2 = 3{,}8$

Los dos sub-problemas obtienen el mismo valor del objetivo que, como era de esperar, es inferior al objetivo inicial Z. Sin embargo, ambos problemas siguen incumpliendo las condición de soluciones enteras. Tenemos que seguir ramificando. Cogemos el problema P1 y lo subdividimos en dos nuevos sub-problemas P11 y P12, añadiendo las restricción $X_2 \leq 4$ en uno y $X_2 \geq 5$ (manteniendo todas las restricciones anteriores, incluida $X_1 \leq 2$). Los resultados son los siguientes:

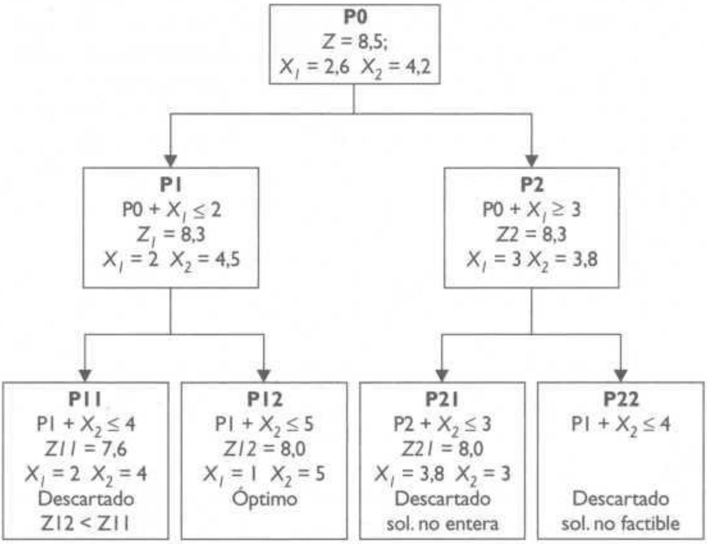
P11 = P1 + $X_2 \leq 4$. Solución: Z11 = 7,6; $X_1 = 2$ y $X_2 = 4$
P12 = P1 + $X_2 \geq 5$. Solución: Z12 = 8,0; $X_1 = 1$ y $X_2 = 5$

En estos dos sub-problemas hemos encontrado soluciones enteras. Como Z11 es inferior a Z12 podemos descartar P11 como solución válida. Por ahora ya hemos encontrado una solución que tiene valores enteros con el problema P12, cuyo objetivo es igual a 8,0. Sin embargo, aún no hemos acabado el algoritmo. Recordemos que habíamos subdividido el problema original en dos sub-problemas. Aún no hemos explorado el segundo sub-problema P2. El valor del objetivo al solucionar P2 era 8,3, aunque la variable X_2 seguía sin ofrecer un valor entero. Como estamos maximizando, podría ser que ramificando P2 en dos sub-problemas se encontrara una solución entera superior a la que hemos encontrado con el sub-problema P12. La ramificación es la siguiente:

P21 = P2 + $X_2 \leq 3$. Solución: Z21 = 8,0; $X_1 = 3{,}8$ y $X_2 = 3$
P22 = P2 + $X_2 \geq 4$. Sin solución.

El problema P22 queda descartado por no tener una solución factible. Los problemas que aún están activos son P12 y P21 y ambos tienen el mismo valor del objetivo. Pero mientras que P12 tiene soluciones enteras, P21 sigue con soluciones fraccionadas. Por lo tanto, podemos descartar P21, ya que, si ramificáramos este problema, al añadir una nueva restricción el valor del objetivo sería inferior (o igual), pero nunca superior. En otras palabras, nunca podríamos encontrar una solución mejor que la que tenemos con P12. Como P12 es la única rama activa, ya tenemos la solución óptima de nuestro problema. Si P21 hubiera dado un valor del objetivo superior a 8,0 con alguna solución fraccionada, tendríamos que seguir bifurcando este sub-problema. El flujo del algoritmo se muestra en la Figura 4.1.

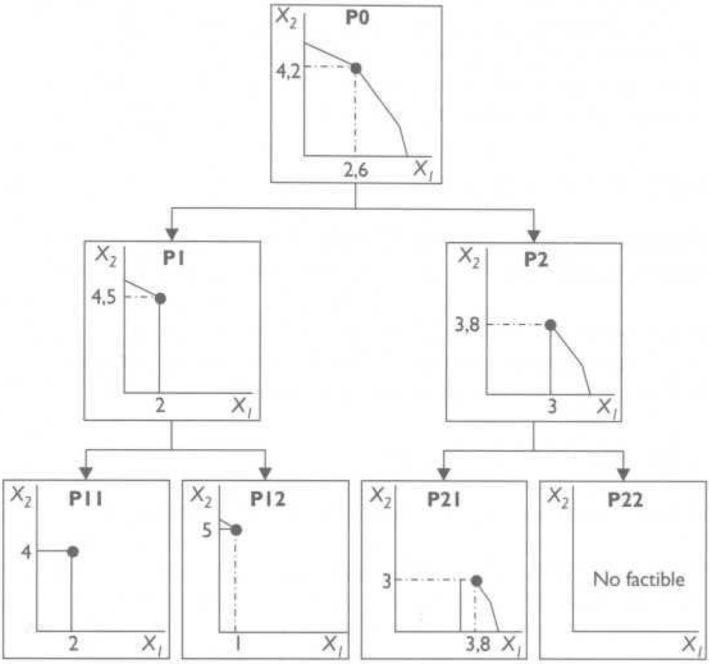
FIGURA 4.1
Árbol del ABA aplicado al ejemplo



En realidad, con este proceso lo que se está haciendo es seccionar en cada ramificación el espacio de soluciones para explorar si existe una solución entera. Este proceso puede ser observado en la Figura 4.2, en donde, para cada sub-problema, se muestra el segmento del espacio de soluciones explorado.

El algoritmo de bifurcación y acotamiento también se utiliza para resolver los problemas de programación entera binaria. La única diferencia es que las variables están acotadas por 0 y 1. Si en un

FIGURA 4.2
Espacio de soluciones de los sub-problemas



sub-problema una variable teóricamente binaria X es igual a 0,6, éste se subdivide en dos problemas: el primero añadirá la restricción X = 0 y el segundo la restricción X = 1.

4.3. Programación entera binaria I: el problema de la mochila

El problema de la mochila es un clásico de la investigación operativa. En esencia, el problema consiste en llenar una mochila con objetos con pesos diferentes y con valores también diferentes. El objetivo es la maximización del valor total de la mochila con la restricción de que el peso de ésta no puede sobrepasar un límite predeterminado. Este problema ha sido utilizado en muchas aplicaciones diferentes. Una de ellas consiste en la asignación de pacientes a una unidad (un ambulatorio, un quirófano, etc.) que tiene una capacidad límite. Cada paciente tiene asociado dos parámetros: el primero mide la gravedad del paciente respecto a los otros en términos relativos, y el segundo mide el tiempo de utilización del servicio. El problema consiste en encontrar qué pacientes podrán ser atendidos y cuáles habrá que derivar a otro centro. Es evidente que en este problema el número de pacientes a ser atendidos es más elevado que la capacidad del centro. En este tipo de problemas nos encontramos con la disyuntiva de que tenemos que, por un lado, dar prioridades para intentar atender el máximo de pacientes, y por el otro lado atender a aquellos que presentan más gravedad.

Formulación del problema

Supongamos que tenemos m pacientes a ser programados en un centro. Los parámetros que tenemos que conocer *a priori* son:

- g_i = valor cardinal de la gravedad del paciente i.
- t_i = duración en minutos de la intervención del paciente i.
- T = tiempo total disponible en el centro.

Y las variables de decisión son:

- X_i = 1, si se atiende al paciente i; 0, si no se le atiende.

En este caso todas la variables son binarias y tendremos una para cada paciente.

Una vez definidos los parámetros y las variables, podemos construir un modelo. Este modelo tiene una única restricción que, básicamente, define que el total de minutos que los pacientes atendidos en el centro consumirán no puede exceder el tiempo total disponible T. Como X_i sólo puede ser igual a 0 ó 1, el tiempo total consumido por el paciente i será igual a t_iX_i. Si sumamos t_iX_i para todos los pacientes, tendremos el tiempo total consumido por ellos, que no puede ser superior a T. En términos matemáticos:

$$\sum_{i=1}^{m} t_i X_i \leq T$$

El objetivo consiste en la maximización de la «gravedad total» del sistema. En otras palabras, queremos atender a aquellos pacientes más necesitados. Tenemos que observar que el problema no es tan trivial, ya que no vale ordenar los pacientes en función de la gravedad e ir llenando «la mochila» del centro hasta agotar la capacidad, porque un paciente j que presenta un nivel de gravedad g_j puede tener asociado un tiempo de atención t_j muy superior al tiempo conjunto de dos pacientes k y l ($t_j > t_k + t_l$), que tienen una menor gravedad, pero cuya gravedad conjunta es superior a la del primero ($g_j < g_k + g_l$). En este caso, sería mejor incluir a los dos pacientes k y l y no al paciente j. El objetivo vendrá definido por la función lineal siguiente:

$$\text{Max } Z = \sum_{i=1}^{m} g_i X_i$$

En definitiva, la formulación final del modelo será:

$$\begin{aligned} \text{Max } Z = & \sum_{i=1}^{m} g_i X_i \\ \text{s. a.} & \\ & \sum_{i=1}^{m} t_i X_i \leq T \\ X_i = (0,1) & \quad i = 1, \ldots, m \end{aligned}$$

Este problema es fácil de resolver utilizando el algoritmo ABA, ya que únicamente tiene una restricción.

Supongamos ahora que algunos tratamientos son incompatibles con otros. Por ejemplo, si se trata de un quirófano, podríamos tener que si operamos al paciente i también podremos operar al paciente j porque tendremos recursos disponibles (quirófano preparado, personal adecuado), pero si no se opera a ningún paciente de tipo i entonces no se podrá operar al paciente j. Para poder introducir esta consideración tendríamos que añadir la siguiente restricción:

$$X_i \geq X_j$$

Es decir, si operamos al paciente i, $X_i = 1$, lo que implica que X_j quedará libre para coger el valor 0 ó 1 (será el modelo quien lo decida). Por otro lado, si $X_i = 0$, la variable X_j será siempre igual 0, y por lo tanto el paciente j no podrá ser operado.

Supongamos ahora que el paciente j únicamente podrá ser operado si tanto el paciente i como el k son operados. En cualquier otro caso el paciente j no podrá ser atendido. Para formular este tipo de restricciones a veces es muy útil la utilización de la «tabla de la verdad». En esta tabla se introduce las combinaciones de las X que son factibles. Esta tabla se representa en el Cuadro 4.1.

La ecuación de esta restricción que tendrá que añadirse al modelo es:

$$X_i + X_k \geq 2\, X_j$$

CUADRO 4.1
Tabla de la verdad

X_i	X_k	X_j
0	0	0
1	0	0
0	1	0
1	1	1

Si X_i y X_k son ambas iguales a 1, X_j podrá ser igual a 0 ó 1. En cualquier otro caso, X_j siempre será igual a 0.

Como hemos visto en este ejemplo, el uso de variables binarias puede ser muy útil para modelizar situaciones en donde la decisión es «hacer o no hacer».

4.4. Programación entera binaria II: problemas de localización de servicios

¿Cuántas ambulancias se necesitan en un área geográfica y dónde deberían ubicarse para asegurar un buen servicio a las llamadas por urgencias? ¿En dónde deberían localizarse los Centros de atención Primaria en una región para minimizar el tiempo de desplazamiento de los usuarios? ¿En dónde tenemos que localizar almacenes para optimizar la distribución de productos farmacéuticos en un país? Estas cuestiones, relacionadas con el diseño y la operación de los servicios de atención y de distribución, han sido estudiadas durante los últimos veinticinco años por un gran número de investigadores. Los planificadores tienen que responder a preguntas como éstas cuando se enfrentan al diseño o a la reconfiguración de los servicios de urgencias médicas, de ambulatorios, de operaciones de distribución o de redes hospitalarias.

Por ejemplo, la velocidad de reacción de un sistema de emergencia a una llamada es el criterio principal para juzgar el desempeño de los servicios de emergencia. Otra medida es la habilidad del personal para lidiar efectivamente con la situación una vez llegado a la escena. La localización inicial de los servidores (parques de bomberos, garajes de ambulancias, etc.) influencia poderosamente la eficiencia de la respuesta. Esto se refleja en la gran cantidad de modelos desarrollados para ayudar a los planificadores de servicios de urgencias.

El problema básico trata de localizar servicios que van a permanecer en su ubicación por un largo tiempo una vez decidida su localización. En otras palabras, su ubicación será, si no definitiva, constante durante un largo período de tiempo. La localización de estos servicios puede ser determinante en la evaluación de la eficiencia de su «desempeño» en la oferta del servicio en cuestión.

Efectivamente, el auge de la investigación operativa en los años sesenta provocó la aparición de un campo específico dedicado a la localización de servicios en regiones y en zonas urbanas. En general, estos modelos optimizan uno o varios objetivos en función de unos recursos limitados y/o criterios de cobertura y de atención. Estos modelos se pueden agrupar en tres categorías en función del objetivo principal. La primera categoría corresponde a modelos cuyo objetivo principal es la maximización de la cobertura de la población siguiendo un criterio «estándar» (por ejemplo, maximizar la población cubierta por el servicio de ambulancias en un tiempo máximo de 10 minutos). El segundo grupo corresponde a modelos de localización cuyo objetivo es la minimización de la distancia o tiempo medio de acceso a la población. El tercero consiste en la minimización de los costes de transporte de mercancías o de personas y de localización de centros.

Entre estos modelos se han realizado diversas variaciones para intentar reflejar algunos aspectos específicos del problema de localización. Por ejemplo, hay modelos que no tan sólo localizan ambulancias, sino que también determinan para cada estación cuál es la combinación óptima de vehículos, materiales y recursos humanos. Otros modelos estudian el problema de la localización teniendo en cuenta el grado de congestión del servicio, intentando obtener un conjunto de localizaciones que no tan sólo optimice la cobertura, sino que de alguna forma considere la situación de que un servicio está ocupado atendiendo una llamada y en su estación se produzca otra llamada (modelos de *backup* o servicios auxiliares). Otra línea de trabajo estudia la situación en donde la demanda de servicio tiene elementos probabilísticos en función de la hora del día.

En esta sección formularemos tres modelos básicos de localización de servicios.

4.4.1. *Modelos de cobertura*

Los modelos de cobertura suelen fijar una distancia estándar D entre un servicio y la población usuaria. Esta distancia (o tiempo de desplazamiento) se considera como la distancia máxima entre usuario y servicio para ofrecer una atención correcta. Esta distancia estándar se utiliza como criterio básico para obtener la ubicación óptima de servicios.

El Problema de localización de servicios con cobertura [17]
El Problema de Localización de Servicios con Cobertura (PLSC), en palabras, es el siguiente:

> ¿Cuál es el número mínimo de centros y dónde tenemos que localizarlos para que toda la población esté cubierta dentro de la distancia estándar D?

El PLSC puede ser formulado de la siguiente forma. Supongamos una red de transporte en donde existen m nodos o áreas, cada uno con una demanda (o población) determinada del servicio, y conectados entre ellos por «arcos» (carreteras, calles, etc.), cada uno de ellos con un tiempo de desplazamiento o una distancia asociados. La localización de los servicios se realiza exclusivamente en los nodos. En otras palabras, únicamente los nodos de la red son candidatos a obtener una localización de los servicios, y por otro lado, los nodos también representan los centros de demanda de los servicios. Muchas veces, en algunas aplicaciones, no todos los nodos de demanda son candidatos a obtener un centro de servicio (a veces en un nodo se encuentra un edificio de interés cultural, o simplemente no existe terreno disponible para ubicar un servicio); por ello siempre en la formulación del modelo se diferencia entre la demanda (nodos que requieren del servicio) y la oferta (nodos candidatos a recibir el servicio). En la Figura 4.3 se representa ejemplo de red de 20 nodos, con la cual formularemos algunos modelos de localización.

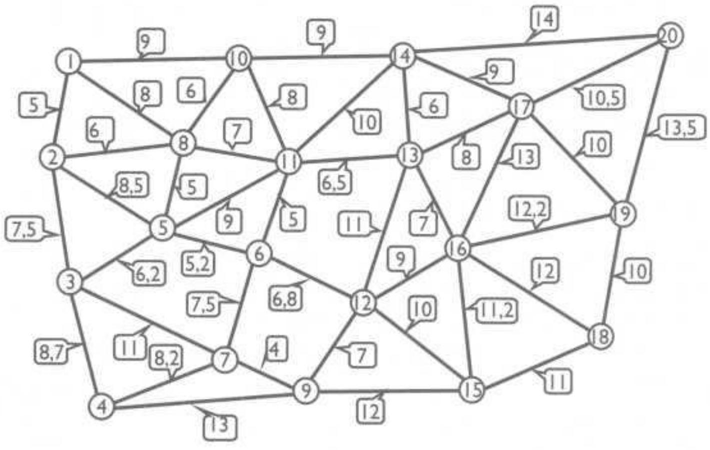
En esta Figura, los nodos están representados por círculos y también se indican las distancias entre los nodos que están directamente conectados. El Cuadro A5.1 del anexo 5.A. contiene la matriz de distancias d_{ij} entre todos los pares de nodos (i,j) de la red. Esta matriz es fundamental en los modelos de localización. En general, esta matriz se obtiene calculando, para cada par de nodos, el «camino más corto» entre ellos, es decir, la distancia más corta que los une.

Ahora es necesario conocer, para cada nodo de demanda, cuáles son las ubicaciones potenciales donde, si se abre un centro en ellas, el nodo de demanda estará cubierto dentro de la distancia estándar D. Por ejemplo, si D = 10, el centro de demanda 4 tiene, como ubicaciones potenciales de cubrirlo, los nodos 3, 4 y 7. En el Cuadro 4.2. se indican, para cada nodo, las ubicaciones potenciales de cobertura.

---

[17] En inglés: «Location Set Covering Problem».

FIGURA 4.3
Red de 20 nodos



CUADRO 4.2
Ubicaciones potenciales de cobertura por nodo. D = 10

Nodo a cubrir	Ubicaciones potenciales	Nodo a cubrir	Ubicaciones potenciales
1	1,2,8,10	11	5,6,8,10,11,13,14
2	1,2,3,5,8	12	6,9,12,15,16
3	2,3,4,5	13	11,13,14,16,17
4	3,4,17	14	10,11,13,14,17
5	2,3,5,6,8,11	15	12,15
6	5,6,7,11,12	16	12,13,16
7	4,6,7,9	17	13,14,17,19
8	1,2,5,8,10,11	18	18,19
9	7,9,12	19	17,18,19
10	1,8,10,11,14	20	20

Por ejemplo, si ubicamos un centro en el nodo 8, los nodos 1, 2, 5, 8, 10 y 11 estarán cubiertos, ya que él está incluido en el conjunto de ubicaciones potenciales de cada uno de estos nodos de demanda.

Definamos X_j como una variable binaria (0 ó 1) que, si es igual a 1, indicará que estamos abriendo un centro en el nodo j, y que, en caso contrario (igual a 0), el nodo j estará vacío. Tendremos tantas variables como ubicaciones potenciales. Podemos utilizar estas variables para formular el modelo. Por ejemplo, tenemos que, para el nodo 4, podemos escribir la siguiente restricción:

$$X_3 + X_4 + X_7 \geq 1$$

que indica que como mínimo una de las tres variables tiene que ser igual a 1, o, en otras palabras, que para que el nodo 4 esté cubierto tenemos que abrir como mínimo un centro en 3, 4, ó 7.

Si definimos N_i como el conjunto de ubicaciones potenciales que cubrirán el nodo i dentro de la distancia estándar D, ($N_i = \{j / d_{ij} \leq D\}$, para cada nodo demanda i podemos escribir la restricción siguiente:

$$\sum_{j=N_i} X_j \geq 1 \qquad i=1,\dots,20$$

Este conjunto de restricciones forzarán a que cada nodo esté cubierto. Ahora falta formular el objetivo. Como queremos minimizar el número de centros a ubicar, cuantas menos X_j sean igual a 1, mejor. Por lo tanto, el objetivo lo podemos formular de la siguiente forma:

$$\text{Min } Z = \sum_{j=1}^{20} X_j$$

Si definimos m como el número total de nodos de demanda y n como el número de total de ubicaciones potenciales, la formulación final del Problema de Localización con Cobertura es:

$$\text{Min } Z = \sum_{j=1}^{n} X_j$$

s.a.

$$\sum_{j=N_i} X_j \geq 1 \qquad i=1,\dots,m$$

$$X_j = (0, 1) \qquad j=1,\dots,n$$

En el archivo mlcp.xls está formulado el problema para nuestro ejemplo, y también se incluye la matriz de distancias. Se puede utilizar el módulo Solver de Excel para resolver el problema con distancias estándar diferentes. En el Cuadro 4.3 se presentan algunos resultados para coberturas diferentes.

Este problema suele tener a veces bastantes soluciones óptimas alternativas. El PLSC puede modificarse para considerar, por ejemplo, la minimización del presupuesto. Si cada nodo potencial de obtener un servicio tiene asociado un coste fijo de apertura f_j, podemos reformular el objetivo del problema de la siguiente forma:

CUADRO 4.3
Resultados con diferentes coberturas

Distancia estándar D	Número mínimo de Centros	Ubicaciones finales
15	3	8,9,19
12	4	3,8,15,17
11	5	3,10,12,18,20
10	6	4,8,12,17,18,20
9	8	4,8,12,15,17,18,19,20
8	9	4,5,8,9,13,15,18,19,20

$$\text{Min } Z = \sum_{j=1}^{20} f_j X_j$$

En este caso estamos minimizando el coste total de apertura de centros. Es muy improbable que aparezcan soluciones alternativas óptimas.

En este problema, la fijación de la distancia estándar D es determinante de los resultados, por lo que hay que ir con mucho cuidado al determinarla.

Supongamos que, en nuestro ejemplo, la distancia estándar está fijada en 10. Al resolver el PLSC encontramos que la solución óptima es igual a 6 centros, ubicados en los nodos 4, 8, 12, 17, 18 y 20. Pero el sistema no tiene suficiente presupuesto para construir 6 centros. Únicamente tiene un presupuesto para 4 centros. Como el número mínimo de centros para cubrir la población es igual a 6, con 4 centros no podremos cubrirla por completo. En este caso, si fijamos en número de centros, podemos intentar encontrar sus ubicaciones de forma a maximizar la cobertura de la población. Este problema es conocido como el Problema de Localización con Cobertura Máxima (PLCM).

El Problema de localización con cobertura máxima (PLCM)
Este problema puede ser descrito de la siguiente forma:

> ¿Dónde tenemos que localizar p centros para maximizar la cobertura de la población dentro de la distancia estándar D?

Este problema es una extensión del PLSC. Para formular el modelo tenemos que añadir un nuevo grupo de variables binarias Y_i, que denominaremos de cobertura, que serán igual a 1 si el nodo i está cubierto por un centro dentro de la distancia estándar D; e igual a

0 si no lo está. Como en la solución final algunos nodos de demanda quedarán descubiertos (ya que no tenemos suficientes centros para cubrir toda la población), algunas de estas variables serán 0. Cojamos de nuevo como ejemplo el nodo 4. Para que él esté cubierto se tiene que localizar como mínimo un centro en uno de los nodos 3, 4 ó 7. Ahora tendremos que escribir la restricción siguiente:

$$X_3 + X_4 + X_7 \geq Y_4$$

Si como mínimo una de las variables de localización X es igual a 1, la variable Y_4 podrá ser también igual a 1, y por lo tanto el nodo de demanda 4 estará cubierto. Si, en cambio, todas las variables X de la restricción son iguales a 0, la variable de cobertura Y_4 será forzosamente igual a 0 y el nodo 4 no estará cubierto. Para cada nodo de demanda escribiremos la siguiente restricción:

$$\sum_{j=N_i} X_j \geq Y_i \qquad i=1,\ldots,20$$

Otra restricción es la limitación del número de centros a localizar. En nuestro ejemplo hemos fijado el número de centros en cuatro. Esto quiere decir que únicamente cuatro variables de ubicación X_j podrán ser igual a 1. La restricción, en términos matemáticos, es:

$$\sum_{j=1}^{20} X_j = 4$$

Finalmente, el objetivo consiste en la maximización de la cobertura de la población de la región en cuestión. En el modelo PLSC todos los nodos quedaban cubiertos, por lo que no hacía falta preocuparse de la población. En el nuevo modelo, como algunos nodos de demanda quedarán descubiertos, tenemos que considerar la población de cada uno de ellos. El modelo intentará cubrir, en primer lugar, aquellos nodos con mayor población (o demanda). El objetivo se formula matemáticamente de la forma siguiente:

$$\text{Max } Z = \sum_{i=1}^{20} a_i Y_i$$

En donde a_i es un parámetro que denota el volumen de demanda (en nuestro ejemplo, población) asociada al nodo i. Contra más Y_i sean igual a 1, más cobertura obtendremos. La formulación final del problema es:

$$\text{Max } Z = \sum_{i=1}^{m} a_i Y_i$$

s. a.

$$\sum_{j=N_i} X_j \geq Y_i \qquad i = 1,\ldots,m$$

$$\sum_{j=1}^{n} X_j = p$$

$$X_j, Y_i = (0,1) \quad j = 1,\ldots,n \quad i = 1,\ldots,m$$

En el archivo plcm.xls está formulado el problema de máxima cobertura para nuestro ejemplo, con D = 10 y p = 4. Las poblaciones de cada uno de los nodos se indica en el Cuadro 5A.1 del anexo. El resultado final se presenta en el Cuadro 4.4.

Cuadro 4.4
Resultados del PLCM en el ejemplo

Número de centros Distancia estándar	4 10
Población cubierta Ubicaciones Nodos no cubiertos	88 % 4,8,12,17 18,20

Vemos que con cuatro centros pasamos a cubrir el 88 % de la población. En otras palabras, mientras que el número mínimo de centros necesarios para cubrir toda la población era igual a 6, con 4 centros cubrimos el 88 % de ella y únicamente dos nodos no están cubiertos.

4.4.2. *Modelo de localización P-mediano*

El modelo P-mediano de localización (MPML) tiene como objetivo principal la minimización de la distancia media entre los nodos de demanda y los centros. El problema, en palabras, es el siguiente:

> ¿Dónde se ubicarán p centros de forma a minimizar la distancia media entre éstos y los nodos de demanda?

Para formular este problema necesitamos conocer, como en el problema anterior, la matriz de distancias y la demanda que se genera en cada uno de los nodos. En este caso no se utiliza una distancia estándar. Las variables de modelo serán:

- $X_{ij} = 1$, si el nodo de demanda i es atendido por el centro ubicado en j; 0, en caso contrario.
- $W_j = 1$; si ubicamos un centro en j; 0, en caso contrario.

A continuación definimos las restricciones. En primer lugar, un nodo de demanda tiene que estar asignado a un único centro. Para forzar esta situación, para cada nodo de demanda i, la suma de las X_{ij} con respecto al índice j tiene que ser igual a 1. En términos matemáticos:

$$\sum_{j=1}^{n} X_{ij} = 1 \qquad i = 1, \ldots, m$$

Ahora bien, el nodo i no podrá ser asignado al nodo j si no existe un centro en j. Como la variable W_j indica si existe un centro en j o no, tendremos que:

$$X_{ij} \leq W_j \qquad i = 1, \ldots, m \qquad j = 1, \ldots, n$$

Si no existe ningún centro en j, $W_j = 0$, lo que implica que ningún nodo de demanda podrá ser asignado a j. En este caso, todas las variables X_{ij} serán igual a 0.

Finalmente, tenemos que fijar el número de centros a abrir. La siguiente restricción tiene que ser añadida al modelo:

$$\sum_{j=1}^{n} W_j = p$$

Finalmente, tenemos que formular el objetivo de distancia media. Tenemos que $a_i d_{ij}$ será la distancia total entre la población (o demanda) en i y el centro en j. Si sumamos $a_i d_{ij}$ para todas las i tendremos toda la demanda asignada a j. Luego tenemos que sumar para todas las j, y tendremos la distancia total del sistema. El objetivo es:

$$\text{Min } Z = \sum_{i=1}^{m} \sum_{j=1}^{n} a_i d_{ij} X_{ij}$$

Una vez obtenido el valor de Z, tenemos que dividirlo por la demanda (o población) total del sistema para obtener la distancia media entre la demanda y los centros.

En resumen, la formulación del problema P-mediano es:

$$\text{Min } Z = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i d_{ij} X_{ij}$$

$$\text{s. a.}$$

$$\sum_{j=1}^{n} X_{ij} = 1 \qquad i=1,\dots,m$$

$$X_{ij} \leq W_j \qquad i=1,\dots,m \quad j=1,\dots,m$$

$$\sum_{j=1}^{n} W_j = p$$

Esta formulación suele tener muchas variables y restricciones. Por ejemplo, si m = n = 100 tendremos 10.100 variables y 10.101 restricciones. Existen varias formas de reducir el número de variables y restricciones, pero aún así este modelo suele ser bastante grande. Se han desarrollado varios métodos heurísticos para poder encontrar soluciones del MPML. Una excelente referencia se encuentra en el libro de Daskin (1995).

4.4.3. *El problema de localización de plantas con capacidad*

En muchos casos el objetivo principal es encontrar una serie de ubicaciones que minimicen tanto los costes de transporte como el coste de apertura de los centros. El modelo de Localización de Plantas con Capacidad (MLPC) se describe de la forma siguiente:

> ¿Cuántos centros se necesitan y dónde hay que ubicarlos para minimizar los costes totales del servicio sin exceder su capacidad?

Para poder formular el modelo, necesitamos los siguientes parámetros:

- a_i = demanda en el nodo i.
- d_{ij} = distancia entre el nodo de demanda i y el nodo de ubicación potencial j.
- f_j = coste de apertura de un centro en el nodo j.
- c_{ij} = coste de transporte por unidad de demanda y unidad de distancia.
- C_j = capacidad de un centro si se ubica en j.

y las variables que a continuación se describen:

- X_{ij} = demanda del nodo i atendida por el centro en j.
- W_j = 1, si ubicamos un centro en j; 0, en caso contrario.

Un vez definidos los parámetros y las variables del modelo, se tienen que formular las restricciones. En primer lugar, no podemos exceder la capacidad de cada centro. Para que esto se cumpla, la demanda asignada a cada uno de los centros potenciales j no puede exceder su capacidad. En términos matemáticos:

$$\sum_{i=1}^{m} X_{ij} \leq C_j \qquad j=1,\ldots,n$$

Por otro lado, la demanda de cada uno de los nodos tiene que ser atendida. Es decir:

$$\sum_{j=1}^{n} X_{ij} = a_i \qquad i=1,\ldots,m$$

Finalmente, un nodo de demanda i no puede ser servido por j si no existe un centro en j. Matemáticamente,

$$\sum_{i=1}^{m} X_{ij} \leq MW_j \qquad j=1,\ldots,n$$

en donde M es un parámetro con un valor muy elevado (por ejemplo, 10^{10}). Si W_j es igual a 0, forzosamente todas las X_{ij} serán igual a 0. En caso contrario, si W_j es igual a 1, las variables X_{ij} podrán tomar cualquier valor, ya que no tendrán ninguna cota superior.

Ahora falta definir el objetivo. Por un lado tenemos los costes de apertura, o costes fijos. Por otro lado, tenemos que minimizar los costes de distribución o transporte. Estos dos tipos de costes juegan un papel opuesto en relación al número de centros a ubicar. Mientras que, contra más centros abramos, menor será el coste de transporte al reducirse las distancias, por otro lado los costes de apertura aumentarán considerablemente. El modelo buscará el número de centros que minimice los costes totales. El objetivo se define matemáticamente como:

$$\text{Min } Z = \sum_{j=1}^{n} f_j W_j + \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} d_{ij} X_{ij}$$

El primer término del lado derecho de la ecuación refleja los costes de apertura. El segundo término formula los costes totales de transporte.

En resumen, el modelo se formula de la siguiente forma:

$$
\begin{aligned}
\text{Min } Z = \; & \sum_{j=1}^{n} f_j W_j + \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} d_{ij} X_{ij} \\
\text{s. a.} \; & \\
& \sum_{i=1}^{m} X_{ij} \le C_{ij} && j=1,\dots,n \\
& \sum_{j=1}^{n} X_{ij} = a_i && i=1,\dots,m \\
& \sum_{i=1}^{m} X_{ij} \le MW_j && j=1,\dots,n \\
& X_{ij} \ge 0; \quad W_j = (0,1) && i=1,\dots,m \quad j=1,\dots,n
\end{aligned}
$$

Este problema es de programación lineal entera mixta, ya que mientras que algunas de las variables son enteras (en este caso las W_j son binarias), otras son continuas (las X_{ij} pueden tomar cualquier valor no-negativo).

Este problema es muy utilizado para ubicar plantas de producción y depósitos de distribución. Existe un sinfín de problemas de localización basados en este modelo.

4.5. Conclusiones

En este capítulo hemos examinado cómo formular algunos problemas de programación entera mixta y binaria y cómo resolverlos con el algoritmo de bifurcación y acotamiento. Mientras que algunos problemas son relativamente fáciles de resolver con este algoritmo, otros pueden ser extremadamente «caros» en términos de tiempo de ordenador, ya que la ramificación es exponencial, y en cada rama del árbol del algoritmo tenemos que resolver un programa lineal. La mayoría de programas lineales incorporan un módulo de programación entera, por lo que no hay que realizar el algoritmo; el propio programa se encarga del proceso. En la hoja de cálculo Excel, para resolver programas con algunas o todas las variables enteras, basta declararlas como enteras o binarias (si procede) dentro de la ventana de restricciones (ver archivos de ejemplo plcm.xls y plsc.xls).

4.6. Problemas

4.1. Resolver el problema siguiente con el algoritmo de bifurcación y acotamiento:

$$\begin{aligned} \text{Max } Z = \ & X_1 + 2X_2 \\ \text{s. a.} \ & \\ & 2{,}1X_1 + 8X_2 \leq 15{,}5 \\ & X_1 + 1{,}2X_2 \leq 5{,}6 \\ & X_1, X_2 \geq 0 \end{aligned}$$

Programar el problema en la hoja de cálculo Excel y resolverlo.

4.2. Resolver el problema siguiente con el algoritmo de bifurcación y acotamiento. En cada iteración del algoritmo, utilizar la hoja de cálculo Excel para resolver el sub-problema correspondiente.

$$\begin{aligned} \text{Max } Z = \ & 4X_1 + X_2 + X_3 \\ \text{s. a.} \ & \\ & X_1 + 4X_2 + X_3 \leq 11{,}5 \\ & 5X_1 - 2X_2 + 4X_3 \leq 10{,}2 \\ & 2{,}1X_1 + 3X_2 + 3X_3 \leq 20 \\ & X_1, X_2, X_3 \geq 0 \end{aligned}$$

4.3. Resolver el problema siguiente con el algoritmo de bifurcación y acotamiento. En cada iteración del algoritmo, utilizar la hoja de cálculo Excel para resolver el sub-problema correspondiente.

$$\begin{aligned} \text{Max } Z = \ & X_1 - X_2 + X_3 \\ \text{s. a.} \ & \\ & X_1 + 4X_2 \leq 20 \\ & X_1 + X_2 + X_3 > 10 \\ & -2X_1 + 4X_2 + 2X_3 \leq 60 \\ & 2X_1 + 3X_2 + X_3 \leq 50 \\ & X_1, X_2, X_3 \geq 0 \end{aligned}$$

4.4. Una red de atención sanitaria tiene m ambulatorios y que atienden a n pueblos de una región rural. La gerencia no está satisfecha con la asignación actual entre pueblos y ambulatorios, y quiere reorganizar la asignación de forma a minimizar los costes de desplazamiento totales del sistema, siempre que no se exceda la capacidad de cada uno de los centros. Sean c_{ij} el coste de transporte entre el pueblo i y el ambulatorio j, a_i la población del pueblo i, y a_j la capacidad del ambulatorio j. Formular un modelo de programación entera que asigne cada pueblo a un único ambulatorio sin exceder la capacidad.

4.5. Una ciudad quiere rediseñar su sistema de ambulancias y mejorar el servicio, sobre todo en lo que se refiere a tiempos de atención. Para ello ha decidido reubicar los garajes. Se tiene la siguiente información: la ciudad está dividida en 20 áreas. En el Cuadro A5.1 del anexo se presentan los tiempos de desplazamiento entre todas las áreas y la frecuencia de llamadas por semana en cada una de ellas (fila *pob.)*. Por ahora se dispone de 4 ambulancias ubicadas actualmente en los nodos 5, 7, 16 y 17, y no se pueden comprar más. El alcalde considera que se tendría que maximizar la cobertura en un tiempo de diez minutos, pero también que toda la ciudad tendría que estar cubierta en quince minutos. Formular el problema y resolverlo con ayuda del ordenador. ¿En cuánto ha mejorado la cobertura respecto a la situación inicial?

Suponer ahora que sólo se tiene presupuesto para reubicar dos de las cuatro ambulancias. Reformular el problema para que el modelo escoja cuáles se tienen que reubicar.

4.7. Anexo: Datos de la red de 20 nodos

Cuadro 5A.1
Población de cada nodo y distancias entre nodos

Pob.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	34	65	99	85	95	75	32	34	28	53	52	21	39	98	69	67	22	54	76	90
1	0,0	5,0	12,5	21,2	13,0	18,2	23,5	8,0	27,5	9,0	15,0	25,0	21,5	18,0	35,0	28,5	27,0	40,5	37,0	32,0
2	5,0	0,0	7,5	16,2	8,5	13,7	18,5	6,0	22,5	12,0	13,0	20,5	19,5	21,0	30,5	26,5	27,5	38,5	37,5	35,0
3	12,5	7,5	0,0	8,7	6,2	11,4	11,0	11,2	15,0	17,2	15,2	18,2	21,7	25,2	27,0	27,2	29,7	38,0	39,4	39,2
4	21,2	16,2	8,7	0,0	14,9	15,7	8,2	19,9	12,2	25,9	20,7	19,2	27,2	30,7	24,2	28,2	35,2	35,2	40,4	44,7
5	13,0	8,5	6,2	14,9	0,0	5,2	12,7	5,0	16,7	11,0	9,0	12,0	15,5	19,0	22,0	21,0	23,5	33,0	33,2	33,0
6	18,2	13,7	11,4	15,7	5,2	0,0	7,5	10,2	11,5	13,0	5,0	6,8	11,5	15,0	16,8	15,8	19,5	27,8	28,0	29,0
7	23,5	18,5	11,0	8,2	12,7	7,5	0,0	17,7	4,0	20,5	12,5	11,0	19,0	22,5	16,0	20,0	27,0	27,0	32,2	36,5
8	8,0	6,0	11,2	19,9	5,0	10,2	17,7	0,0	21,7	6,0	7,0	17,0	13,5	15,0	27,0	20,5	21,5	32,5	31,5	29,0
9	27,5	22,5	15,0	12,2	16,7	11,5	4,0	21,7	0,0	24,5	16,5	7,0	18,0	24,0	12,0	16,0	26,0	23,0	28,2	36,5
10	9,0	12,0	17,2	25,9	11,0	13,0	20,5	6,0	24,5	0,0	8,0	23,0	14,5	9,0	33,0	21,5	18,0	33,5	28,0	23,0
11	15,0	13,0	15,2	20,7	9,0	5,0	12,5	7,0	16,5	8,0	0,0	11,8	6,5	10,0	21,8	13,5	14,5	25,5	24,5	24,0
12	25,0	20,5	18,2	19,2	12,0	6,8	11,0	17,0	7,0	23,0	11,8	0,0	11,0	17,0	10,0	9,0	19,0	21,0	21,2	29,5
13	21,5	19,5	21,7	27,2	15,5	11,5	19,0	13,5	18,0	14,5	6,5	11,0	0,0	6,0	18,2	7,0	8,0	19,0	18,0	18,5
14	18,0	21,0	25,2	30,7	19,0	15,0	22,5	15,0	24,0	9,0	10,0	17,0	6,0	0,0	24,2	13,0	9,0	25,0	19,0	14,0
15	35,0	30,5	27,0	24,2	22,0	16,8	16,0	27,0	12,0	33,0	21,8	10,0	18,2	24,2	0,0	11,2	24,2	11,0	21,0	34,5
16	28,5	26,5	27,2	28,2	21,0	15,8	20,0	20,5	16,0	21,5	13,5	9,0	7,0	13,0	11,2	0,0	13,0	12,0	12,2	23,5
17	27,0	27,5	29,7	35,2	23,5	19,5	27,0	21,5	26,0	18,0	14,5	19,0	8,0	9,0	24,2	13,0	0,0	20,0	10,0	10,5
18	40,5	38,5	38,0	35,2	33,0	27,8	27,0	32,5	23,0	33,5	25,5	21,0	19,0	25,0	11,0	12,0	20,0	0,0	10,0	23,5
19	37,0	37,5	39,4	40,4	33,2	28,0	32,2	31,5	28,2	28,0	24,5	21,2	18,0	19,0	21,0	12,2	10,0	10,0	0,0	13,5
20	32,0	35,0	39,2	44,7	33,0	29,0	36,5	29,0	36,5	23,0	24,0	29,5	18,5	14,0	34,5	23,5	10,5	23,5	13,5	0,0

5. PROGRAMACIÓN MULTIOBJETIVO

5.1. Introducción

Como bien cita Romero (1993), el gran economista de la escuela de Chicago Milton Friedman, en el primer capítulo de su libro «Teoría de los Precios» (1962), enfatiza el carácter multicriterio de los problemas económicos al considerar como económicos sólo los problemas en los que subyace la existencia de criterios múltiples. Por el contrario, cuando el problema de decisión se establece en base a un solo criterio, nos encontramos ante lo que Friedman llama un problema tecnológico, en el que no existen problemas de elección propiamente dichos. Efectivamente, en muchas ocasiones, el decisor se enfrenta a situaciones en donde existen varios objetivos a maximizar o minimizar. Por ejemplo, en una organización sanitaria, por un lado se puede querer maximizar una medida del bienestar de la población y por el otro minimizar los costes de implantación del servicio. O una empresa farmacéutica puede querer maximizar beneficios, aumentar cuota de mercado, aumentar la tasa de retorno y minimizar el volumen de *stocks*. Cuando tenemos más de un objetivo en un problema, podemos utilizar la programación multiobjetivo. En este caso, como la optimización simultánea de todos los objetivos es normalmente imposible –ya que en la vida real entre los objetivos que pretende optimizar un centro decisor suele existir un cierto grado de conflicto–, el enfoque multiobjetivo, en vez de intentar determinar un no existente óptimo, pretende establecer el conjunto de soluciones eficientes o Pareto óptimas.

5.2. Espacio de decisiones y espacio de objetivos

Supongamos el programa lineal bicriterio siguiente:

$$
\begin{aligned}
\text{Max } Z_1 &= 2X_1 - X_2 \\
\text{Max } Z_2 &= -X_1 + 5X_2 \\
\text{s. a.} & \\
X_1 + X_2 &\leq 8 \\
-X_1 + 2X_2 &\leq 7 \\
X_1 &\leq 6 \\
X_2 &\leq 4 \\
X_1, X_2 &\geq 0
\end{aligned}
$$

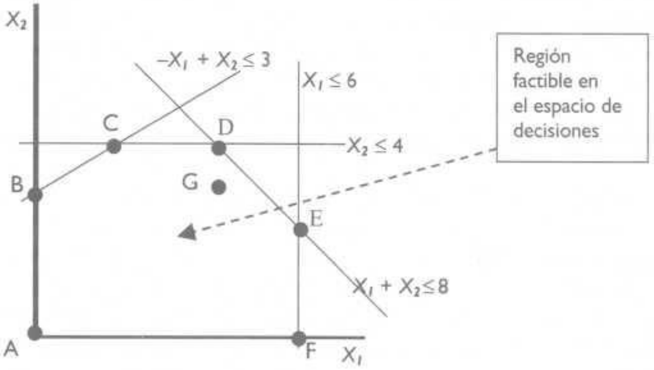
El conjunto de restricciones sigue formando, como en el caso monocriterio, una región factible convexa, que nos indica todas las combinaciones de las variables de decisión X_1 y X_2 que cumplen todas las restricciones del problema. Este conjunto convexo es conocido como el *espacio de decisiones.* Los puntos extremos también pertenecen al espacio de decisiones. Como explicamos en el capítulo 3, el valor óptimo del objetivo siempre corresponderá al valor de las variables en uno de estos puntos extremos. En el cuadro 5.1 se presentan los valores de las variables y de los objetivos en cada uno de los puntos extremos del ejemplo.

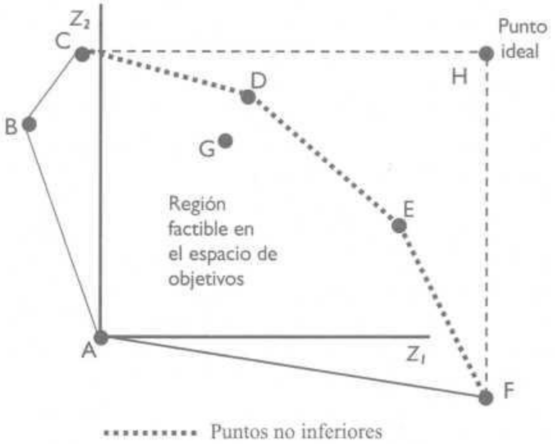
CUADRO 5.1
Valores de las variables y objetivos en los puntos extremos

Puntos extremos	X_1	X_2	Z_1	Z_2
A	0	0	0	0
B	0	3	–3	15
C	1	4	–2	19
D	4	4	4	16
E	6	2	10	4
F	6	0	12	–6

En la Figura 5.1 tenemos dibujado el espacio de decisiones del ejemplo. Los puntos A, B, C D y F corresponden a los puntos extremos del problema, y el punto G es una solución factible. Como para cada valor de X_1 y de X_2 podemos encontrar el valor de los objetivos, podemos también representar en un gráfico los valores de los objetivos que corresponden a cada uno de estos valores de X_1 y X_2, que son factibles para el problema. En la Figura 5.2 se re-

FIGURA 5.1 y 5.2
Espacio de decisiones y Espacio de objetivos





presenta *la región factible en el espacio de objetivos.* Por ejemplo, cada punto extremo en el espacio de decisiones corresponde también a un punto extremo en el espacio de objetivos.

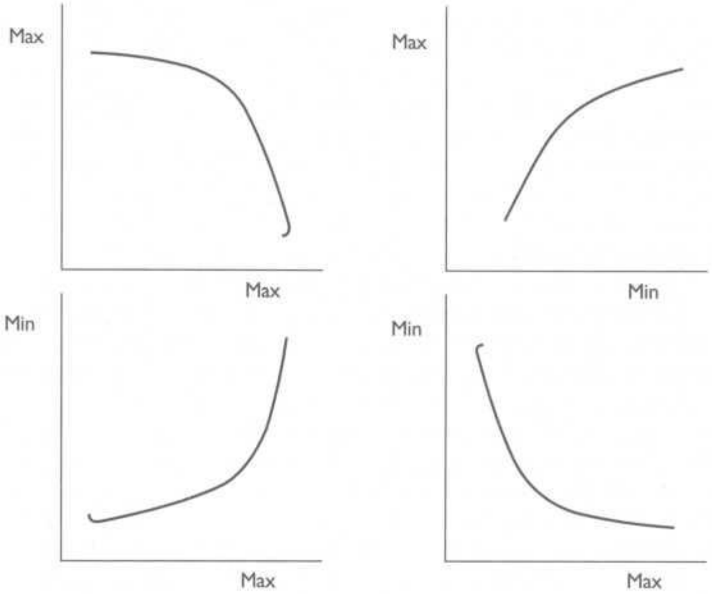
Si estuviéramos únicamente maximizando el primer objetivo, éste sería igual a 10 y $X_1 = 6$ y $X_2 = 2$, que corresponde a punto extremo E. En este punto extremo el valor del segundo objetivo no es óptimo, ya que es igual a 4, mientras que el valor óptimo se alcanza en el punto extremo C, y es igual a 19. Por lo tanto, una solución óptima para uno de los objetivos nos da una solución muy deficiente para el segundo. Pero, por otro lado, tenemos que encontrar una solución única de las variables de decisión que a su vez determinará el valor de los dos objetivos. En este caso tenemos que encontrar una solución *de compromiso* para los dos objetivos. Esta solución puede no dar un resultado óptimo de los objetivos, pero tiene que ser una solución satisfactoria para ambos. Para ello se define el concepto de *no-inferioridad* o *de eficiencia*. Diremos que una solución es *no-inferior* (o *eficiente)* si no existe ninguna otra solución del problema que dé un valor mejor de los objetivos, o que mejore un objetivo, dejando al otro igual.

Observemos las Figuras 5.1 y 5.2. El punto H sería el punto ideal, ya que es donde ambos objetivos alcanzan su valor óptimo. Sin embargo, este punto no es factible. El punto G es factible, pero es un punto inferior porque podemos encontrar otro punto factible que da un mayor valor de ambos objetivos (por ejemplo, el punto extremo D). Por otro lado, el punto D es no-inferior, porque no podemos encontrar otra solución factible del problema sin tener que empeorar uno de los objetivos. Recordemos que estamos maximizando ambos objetivos. Ahora podemos definir el conjunto de puntos no-inferiores. Gráficamente, este conjunto corresponde al borde de la región factible en el espacio de decisiones entre los puntos extremos C-D-E-F. Cualquier valor de Z_1 y de Z_2 situado sobre esta frontera corresponde a un punto no-inferior. Cualquier punto situado en el interior del conjunto será una solución inferior, y por lo tanto no nos interesa considerarla.

Una vez obtenido el conjunto de puntos no-inferiores en el espacio de objetivos, tenemos que escoger uno que nos dé la solución final del problema. Aquí es donde entran las consideraciones subjetivas del decisor, ya que, en muchos casos, las unidades de los objetivos son diferentes, El decisor comparará los diferentes puntos eficientes y los intercambios que se producen al pasar de un punto no-inferior a otro punto. En otras palabras, al pasar del punto D al punto E en nuestro ejemplo, el decisor valorará la diferencia entre el beneficio adicional que obtiene al aumentar Z_1 con la reducción del valor de Z_2. Esta es una de las grandes ventajas de la programación multiobjetivo, ya que ofrece al decisor diferentes soluciones alternativas eficientes. El decisor escogerá aquella que crea mejor y se puede apoyar en otros métodos para poder discriminar entre las diferentes soluciones.

Hasta ahora hemos examinado el problema de maximización de dos objetivos. Pero en muchas situaciones tenemos que maximizar un objetivo y minimizar el otro, o minimizar los dos objetivos. En este caso, la frontera de eficiencia adopta formas diferentes en función de la característica de los objetivos. En la Figura 5.3 se muestran estas formas diferentes.

FIGURA 5.3
Fronteras de eficiencia



5.3. Métodos de resolución

Existen dos métodos clásicos para poder generar las soluciones no-inferiores de un problema multiobjetivo: el método de los pesos y el método de las restricciones. Ambos métodos intentan generar todos los puntos no inferiores del espacio de objetivos, si bien que, una vez aplicados, no garantizan la obtención de todos ellos.

5.3.1. *El método de la restricción*

El método de la restricción consiste básicamente en la transformación del problema multiobjetivo en un problema con un único objetivo a maximizar o minimizar, para poder así utilizar los métodos de resolución clásicos, como el Simplex. En esencia, todos los objetivos del problema, menos uno, se introducen en el conjunto de restricciones, fijando arbitrariamente el lado derecho de cada nueva restricción (una por objetivo). Marglin (1967) demostró que la solución de este nuevo problema también da una solución eficiente. El procedimiento paso a paso es el siguiente para un problema bi-objetivo en donde los dos objetivos son de maximización:

En primer lugar, solucionamos el problema con el primer objetivo Z_1 y, una vez obtenido el valor de las variables de decisión, calculamos el valor del segundo objetivo Z_2. A continuación solucionamos el problema con el segundo objetivo Z_2 y una vez obtenido el valor de las variables de decisión, calculamos el valor del primer objetivo Z_1. Por ahora ya tenemos dos puntos eficientes del espacio de objetivos. El siguiente paso es escoger una de las funciones objetivo (por ejemplo, escogemos Z_2) y ponerla como restricción. Supongamos que L_2 corresponde al valor de Z_2 cuando maximizamos Z_1. Si añadimos la restricción $Z_2 \geq L_2$ y solucionamos el problema, volveríamos a obtener la misma solución para Z_1. Pero si añadimos la siguiente restricción $Z_2 \geq L_2 + \varepsilon$, en donde ε es un valor positivo relativamente pequeño, y solucionamos el problema, es posible que la nueva solución de Z_1 sea inferior o igual, pero obviamente nunca superior, ya que al añadir una nueva restricción estamos reduciendo el espacio de decisiones factible. Por lo tanto, a medida que vamos incrementando el valor de ε y resolviendo nuevas instancias del problema, vamos generando nuevas soluciones de Z_1. El proceso se para cuando el lado derecho de la restricción, $L_2 + \varepsilon$, alcanza el valor óptimo de Z_2. El problema reside en encontrar el valor de ε adecuado para poder generar el máximo número de puntos eficientes en el espacio de objetivos. En otras palabras, se trata de encontrar el número adecuado de problemas lineales a resolver. A continuación vamos a exponer cómo se resuelve el ejemplo multiobjetivo de este cápitulo. El modelo original es el siguiente:

$$
\begin{array}{rrrl}
\text{Max } Z_1 = & 2X_1 & -X_2 & \\
\text{Max } Z_2 = & -X_1 & +5X_2 & \\
\text{s. a.} & & & \\
& X_1 & +X_2 & \leq 8 \\
& -X_1 & +2X_2 & \leq 7 \\
& X_1 & & \leq 6 \\
& & X_2 & \leq 4 \\
& X_1, X_2 \geq 0 & &
\end{array}
$$

Si resolvemos el problema únicamente con el primer objetivo, Z_1, tendremos que la solución de las variables de decisión correspondiente es $X_1 = 6$, $X_2 = 0$, $Z_1 = 12$ y $Z_2 = -6$, que corresponde al punto extremo E. Por el contrario, si resolvemos el problema únicamente con el segundo objetivo, Z_2, tendremos que la solución de las variables de decisión correspondiente será $X_1 = 1$, $X_2 = 4$, $Z_1 = -2$ y $Z_2 = 19$, correspondiente al punto extremo D. Ya tenemos dos puntos extremos eficientes. Ahora se trata de generar otros puntos extremos eficientes del problema. Siguiendo lo expuesto anteriormente, introducimos la ecuación del segundo objetivo como restricción adicional del modelo, con el lado derecho igual a $-6 + \varepsilon$, de la siguiente forma:

$$-X_1 + X_2 \geq -6 + \varepsilon$$

que, multiplicando ambos lados por -1, es equivalente a:

$$X_1 - X_2 \leq 6 - \varepsilon$$

Supongamos ahora que vamos aumentando arbitrariamente el valor de ε en cinco unidades ($\varepsilon = 5,10,15,20$) hasta llegar a $\varepsilon = 25$, ya que en este punto el lado derecho será igual al valor óptimo de Z_2 (19). Para cada valor de ε resolvemos el problema utilizando el método Simplex. En el Cuadro 5.2 se exponen los resultados:

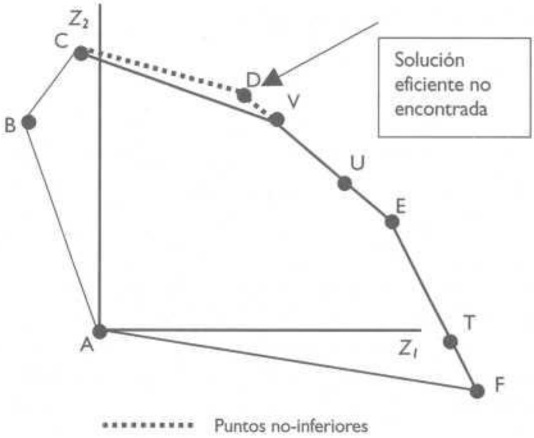
CUADRO 5.2
Resultados del método de la restricción, ejemplo 6.1

ε	$6-\varepsilon$	X_1	X_2	Z_1	Z_2	Puntos de la frontera
0	6	6	0	12	–6	F
5	1	6	1	11	1	T
10	–4	6	2	10	4	E
15	–9	5,7	2,8	7,5	9	U
20	–14	4,3	3,7	5	14	V
25	–19	1	4	–2	19	C

Como podemos ver en esta tabla, a medida que incrementamos el valor de ε, el valor del primer objetivo se va reduciendo debido a la nueva restricción, que cada vez más acota el espacio factible de decisiones. Por otro lado, como es de esperar, el segundo objetivo aumenta al incrementar ε, hasta alcanzar su valor óptimo cuando incluimos la restricción $X_1 - X_2 \leq -25$, que es equivalente a $-X_1 + X_2 \geq 25$.

La pregunta que a continuación nos hacemos es: ¿hemos escogido unos valores aceptables de ε? O, en otras palabras, ¿hemos generado todos los puntos extremos eficientes en el espacio de objetivos? La respuesta, en este caso, es negativa. Esto lo podemos observar en la Figura 5.4. Como podemos ver, hemos generado todos los puntos extremos que conocíamos anteriormente, excepto

FIGURA 5.4
Puntos eficientes del ejemplo 6.1



el punto D. Esto es debido a que el incremento de ε aplicado al lado derecho de la restricción cuando estábamos en el punto V ha sido superior al valor del segundo objetivo en el punto D. Efectivamente, en el punto V, $Z_2 = 14$ y al aumentar ε en cinco unidades, estábamos acotando por debajo en 19 la restricción correspondiente al segundo objetivo. En este caso, nunca podríamos encontrar el valor de Z_2 correspondiente al punto D.

Como hemos visto, el método de la restricción no garantiza la obtención de todos los puntos extremos en el espacio de objetivos. Es importante escoger valores de ε adecuados para poder generar el máximo de puntos eficientes posibles.

Si el objetivo que escogemos para incluirlo en el conjunto de restricciones es de minimización, tendremos que cambiar la dirección de la desigualdad. En este caso tendremos que acotar la ecuación de la siguiente forma: $Z_2 \leq L - \varepsilon$.

5.3.2. *El método de los pesos*

El método de los pesos es otro procedimiento similar al de la restricción para generar los puntos eficientes del espacio de objetivos. También se trata de transformar el programa multiobjetivo en un programa con un único objetivo para poder utilizar el método Simplex y generar así soluciones eficientes. En este caso, básicamente se forma un único objetivo sumando los dos objetivos del modelo ponderados por unos pesos relativos. Para obtener diferentes puntos eficientes estos pesos relativos se van modificando. En cada modificación se resuelve el problema con el nuevo objetivo resultante. Volvamos al ejemplo de la sección anterior. Los objetivos son los siguientes:

$$\text{Max } Z_1 = 2X_1 - X_2$$

$$\text{Max } Z_2 = -X_1 + 5X_2$$

Ahora tenemos que formar un único objetivo. Si definimos w_1 y w_2 como los pesos asociados al primer y segundo objetivo, respectivamente, podemos escribir el nuevo objetivo de la siguiente forma:

$$\text{Max } Z = w_1Z_1 + w_2Z_2$$

Si fijamos $w_1=1$ y $w_2=0$ estaremos resolviendo únicamente el problema con el primer objetivo. Si, al contrario, fijamos $w_1=0$ y $w_2=1$ encontraremos la solución del problema de maximización con el segundo objetivo exclusivamente. Si fijamos valores intermedios, iremos obteniendo soluciones intermedias eficientes. Esto es debido a que, al cambiar los pesos relativos, estamos modificando la pendiente del nuevo objetivo, y podemos ir encontrando puntos extremos diferentes en el espacio de decisiones. Una vez obtenida la solución de las variables de decisión, podemos calcular el valor de cada uno de los objetivos. Como la nueva función objetivo también encuentra su máximo en un punto extremo (propiedad básica que hemos visto en el capítulo de métodos de solución), los puntos obtenidos en el espacio de objetivos también serán eficientes. A continuación resolveremos nuestro ejemplo con el método de los pesos.

En primer lugar formamos la nueva función objetivo:

$$\text{Max } Z = w_1(2X_1 - X_2) + w_2(-X_1 + 5X_2)$$

que es equivalente a:

$$\text{Max } Z = (2w_1 - w_2)\ X_1 + (-w_1 + 5w_2)\ X_2$$

sujeto al conjunto de restricciones original.

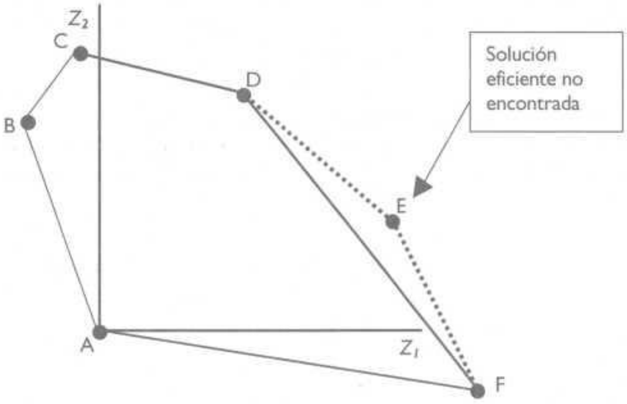
Ahora se trata de ir fijando valores a los pesos arbitrariamente para ir encontrando soluciones diferentes del problema. En el Cuadro 5.5 se presentan los resultados obtenidos. Como podemos

CUADRO 5.5
Resultados del método de la restricción, ejemplo 5.1

w_1	w_2	x_1	x_2	z_1	z_2	Puntos de la frontera
1	0	6	0	12	–6	F
0,80	0,2	6	2	10	4	E
0,60	0,4	4	4	4	16	D
0,4	0,6	4	4	4	16	D
0,2	0,8	1	4	–2	19	C
0	1	1	4	–2	19	C

ver, a medida que aumentamos el valor de w_2 y reducimos el valor de w_1 vamos dando más importancia al segundo objetivo en detrimento del primero. Los valores obtenidos para los objetivos son eficientes, como podemos observar en la Figura 5.5, si bien que no hemos conseguido generar todos los puntos eficientes del espacio de objetivos. Como en el método anterior, el procedimiento de los pesos no garantiza la obtención de todos los puntos eficientes.

FIGURA 5.5
Solución obtenida con el método de los pesos



Un problema adicional de este método es el fijar unos valores relativos de los pesos, debido a que las unidades de los objetivos suelen ser diferentes. Por ejemplo, si por un lado estamos maximizando la cobertura de la población y por el otro maximizamos los beneficios, los pesos relativos que apliquemos tienen que reflejar la relación de estas unidades. Si la cobertura de la población en cuestión está en torno a 10.000 de personas y los beneficios en torno a 10.000.000

de pesetas, tenemos que la relación es de 1 a 1.000. En este caso, el valor del peso asociado al primer objetivo tiene que reflejar esta relación, ya que si no el segundo objetivo siempre tendrá una importancia muy superior y «pesará» siempre mucho más que el primer objetivo cuando utilizamos este método. En este caso fijaríamos valores de w_1 cercanos a 1.000 y de w_2 cercanos a 1.

Si tenemos que Z_1 es de maximización y Z_2 es de minimización, el nuevo objetivo sería el siguiente:

$$\text{Max } Z = w_1Z_1 - w_2Z_2$$

Por el contrario, si ambos objetivos son de minimización, tendremos que minimizar el siguiente objetivo:

$$\text{Min } Z = w_1Z_1 + w_2Z_2$$

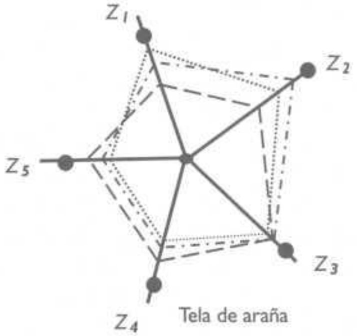
5.3.3. *Extensiones de la programación multiobjetivo*

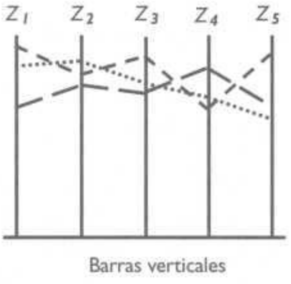
Hasta ahora nos hemos ocupado del caso de problemas con dos objetivos; pero en muchas situaciones podemos tener más de dos objetivos. Los métodos de la restricción y de los pesos siguen siendo válidos, aunque dejan de ser bastante eficientes, ya que el número de combinaciones que tenemos que realizar aumenta exponencialmente. Por ejemplo, si utilizamos el método de los pesos en una formulación con cuatro objetivos, tendremos que definir cuatro pesos y asignarles muchas combinaciones de valores diferentes. Lo mismo sucede con el método de la restricción. Existen otras técnicas, como la programación por metas, que permite definir prioridades entre los diferentes objetivos y así eliminar muchas combinaciones diferentes de problemas a solucionar.

En la programación multiobjetivo es muy importante que exista una gran interacción entre el decisor y el problema. El decisor tiene que escoger una única solución eficiente entre muchas y es por esto que la curva de intercambio (o de eficiencia) es muy útil, porque permite realizar un análisis marginal, preguntándose cuestiones del tipo: si voy del punto D al punto E, ¿me compensa el aumento del primer objetivo la pérdida que pasa en el segundo? Esta interacción puede ser muy útil también en situaciones con múltiples objetivos. El decisor puede fijar *a priori* algunos de los pesos, o dar algún límite inferior o superior para acotar así el campo de soluciones diferentes en el espacio de objetivos.

Otro problema de la programación multiobjetivo es la representación del espacio de decisiones. Como hemos visto en este capítulo, en problemas bi-objetivos esta representación es relativamente sencilla, ya que se puede realizar en un simple gráfico cartesiano. Si pasamos a tener tres objetivos, al querer realizar una representación cartesiana, tendríamos que definir un gráfico de tres dimensiones, que complica bastante el asunto. Con más de tres objetivos, este tipo de representaciones ya no es factible. Sin embargo, existe otro tipo de representaciones que pueden ser útiles para mostrar el espacio de objetivos. En la Figura 5.6 se muestran dos de estos tipos de representaciones para tres soluciones diferentes de un problema con cinco objetivos.

FIGURA 5.6
Representaciones gráficas de soluciones





5.4. Problemas

5.1. Considerar el programa lineal bi-objetivo siguiente:

$$
\begin{aligned}
\text{Max } Z &= -2X_1 + X_2 \\
\text{Max } Z &= 3X_1 - 2X_2 \\
\text{s. a.} & \\
X_1 + 4X_2 &\leq 8 \\
X_1 + X_2 &\leq 5 \\
X_1, X_2 &\geq 0
\end{aligned}
$$

Utilizar el método gráfico para encontrar el espacio de decisiones y el espacio de objetivos. Señalar cuáles son las soluciones eficientes.

5.2. Supongamos el programa lineal bi-criterio siguiente:

$$
\begin{array}{llll}
\text{Max } Z_1 = & X_1 & +2X_2 & -X_3 \\
\text{Max } Z_2 = & -X_1 & +4X_2 & +X_3 \\
\text{s. a.} & & & \\
& X_1 & +4X_2 & +X_3 \leq 12 \\
& 5X_1 & -2X_2 & +4X_3 \leq 11 \\
& 2X_1 & +3X_2 & +3X_3 \leq 20 \\
& & X_1, X_2, X_3 \geq 0 &
\end{array}
$$

Solucionar el problema bi-objetivo con el método de los pesos y encontrar todas las soluciones eficientes.

5.3. Utilizar el método de los pesos para solucionar el problema siguiente:

$$
\begin{array}{llll}
\text{Max } Z_1 = & -X_1 & +X_2 & +2X_3 \\
\text{Min } Z_2 = & X_1 & +5X_2 & + - X_3 \\
\text{s. a.} & & & \\
& 2X_1 & +8X_2 & -X_3 \leq 20 \\
& -X_1 & +X_2 & +X_3 \leq 30 \\
& 2X_1 & +3X_2 & +X_3 \leq 50 \\
& & X_1, X_2, X_3 \geq 0 &
\end{array}
$$

5.4. Considerar el programa lineal bi-objetivo siguiente:

$$
\begin{array}{lll}
\text{Max } Z_1 = & -3X_1 & +2X_2 \\
\text{Max } Z_2 = & X_1 & - 2X_2 \\
\text{s. a.} & & \\
& X_1 & +4X_2 \leq 8 \\
& 3X_1 & +2X_2 \leq 9 \\
& 2X_1 & +X_2 \leq 8 \\
& X_1 & +X_2 \leq 5 \\
& & X_1, X_2 \geq 0
\end{array}
$$

Solucionar el problema con el método de las restricciones.

5.5. Utilizar el método de los pesos para solucionar el problema siguiente:

$$
\begin{array}{lrrrl}
\text{Min } Z = & 2X_1 & +3X_2 & -2X_3 & \\
\text{Min } Z = & 4X_1 & -5X_2 & +7X_3 & \\
\text{s. a.} & & & & \\
& 2X_1 & +8X_2 & -X_3 & \geq 20 \\
& -X_1 & +X_2 & +X_3 & \geq 30 \\
& 2X_1 & +3X_2 & +X_3 & \geq 50 \\
& 7X_1 & -3X_2 & +2X_3 & \geq 54 \\
& 2X_1 & +9X_2 & -X_3 & \geq 37 \\
& & \multicolumn{3}{l}{X_1, X_2, X_3 \geq 0}
\end{array}
$$

6. GESTIÓN DE COLAS

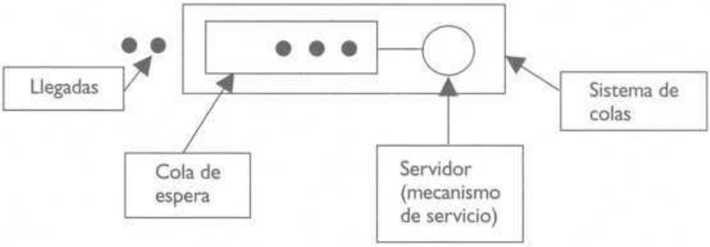
En la mayoría de las organizaciones existen ejemplos de procesos que generan colas de espera. Estas colas suelen aparecer cuando un usuario, un empleado, una máquina o una unidad tienen que esperar a ser servidas debido a que la unidad de servicio, operando a plena capacidad, no puede atender temporalmente a este servicio. Un típico ejemplo de colas de espera que ilustra el problema es un viaje en avión. Primero, para comprar el billete podemos tener que hacer cola en la ventanilla correspondiente. Una vez obtenido el billete, tendremos que hacer cola para facturar el equipaje y obtener las tarjetas de embarque. Después hacemos cola para pasar por el detector de metales y finalmente esperamos en cola en la sala de embarque. Una vez dentro del avión, tendremos que esperar a que los pasajeros coloquen sus bolsas de mano para poder llegar a nuestro asiento. Cuando el avión se dirige hacia la pista de despegue puede encontrarse con una cola de aviones esperando su turno para despegar. Cuando llega a su destino, puede dar unas cuantas vueltas antes de tener permiso para aterrizar. Y finalmente, cuando se asigna una puerta de desembarque para el avión, tendremos que esperar a que lleguen las maletas. En este viaje es posible que hayamos sido miembros de por lo menos diez colas. Y eso sin considerar la experiencia en colas de la propia compañía aérea para este mismo viaje. El avión en el cual viajábamos tiene que esperar en cola para repostar, ser inspeccionado, asignarle una puerta determinada, una tripulación, una carga de comidas, una ruta específica, etc. De ahí que las compañías aéreas se preocupan de gestionar sus operaciones lo más eficientemente posible, y tratar de reducir al mínimo el tiempo de espera en realizar dichas operaciones.

Los sistemas sanitarios también se enfrentan a este tipo de problemas. Las listas de espera son muy comunes en muchos procesos quirúrgicos dentro de una red pública, y a nivel ambulatorial es muy común la existencia de personas esperando a ser atendidas en un Centro de Asistencia Primaria. Los sistemas de urgencias muchas veces se ven congestionados, siendo el tiempo de espera crucial. Los modelos de gestión de colas intentan simular el sistema en donde puede existir congestión (y por lo tanto, colas) y generan una serie de parámetros –que veremos en este capítulo– que permiten evaluar el sistema actual y evaluar la realización de modificaciones en el servicio en cuestión.

6.1. Descripción de un sistema de colas

Un sistema de colas tiene dos componentes básicos: la cola y el mecanismo de servicio. En la figura 6.1 se presenta un esquema de una cola simple.

FIGURA 6.1
Esquema de Cola Simple



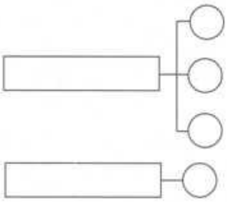
Pueden existir varias configuraciones de colas más complejas. En la figura 6.2 se exponen otros tipos de configuraciones de sistemas de colas.

El proceso básico en la mayoría de los sistemas de colas es el siguiente. Los clientes que vienen a procurar un determinado servicio se generan a través del tiempo en una fuente de entrada. Estos clientes entran dentro del sistema y se unen a una cola. En un determinado momento se selecciona uno de los clientes para poder proporcionarle el servicio en cuestión, mediante lo que se denomina la *disciplina de servicio.* Esta disciplina es la que rige el mecanismo de atención. Una vez seleccionado el cliente, éste es atendido por el *mecanismo de servicio.* Una vez terminado el servicio, el cliente sale del sistema.

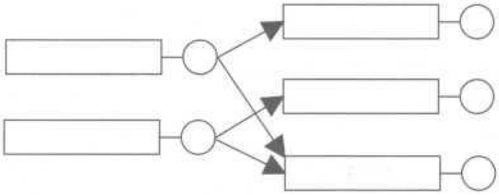
En general, un sistema de colas tiene una *población potencial* infinita. Es decir, que el tamaño de la cola es muy pequeño respecto al potencial de usuarios del sistema. Por ejemplo, un ambulatorio de ur-

FIGURA 6.2
Configuraciones de colas

Sistema de colas en paralelo



Sistema de colas en serie



Sistema complejo de colas

gencias en general cubre una región con población grande comparado con las posibles urgencias que se puedan generar. Ahora bien, existen casos en donde la población es finita respecto del tamaño de la cola. Esto puede suceder en la farmacia de un hospital, en donde la población potencial la forman las enfermeras y ATS. En un momento dado puede formarse una cola considerable. Como los cálculos son mucho más sencillos para el caso infinito, esta suposición se emplea casi siempre.

Otro factor a tener en cuenta es el patrón estadístico mediante el cual se generan los clientes a través del tiempo. La suposición normal es que el proceso se genere siguiendo un *proceso de Poisson,* que veremos más adelante en este capítulo. Si el proceso de llegada es Poisson, el tiempo entre cada una de las llegadas sigue una *distribución exponencial.*

Otro factor importante a tener en cuenta en un sistema de colas es la *«fuga»* de algún cliente. Al modelizar la cola hay que considerar si una persona que lleva dentro de la cola un rato, desiste de ser atendida, cansada de esperar, abandonando la cola.

Como hemos mencionado anteriormente, la disciplina de la cola rige el sistema de entrada en el mecanismo de servicio. La mayoría de los sistemas utiliza el método «First In First Out», conocido como FIFO. Otros sistemas pueden ser de tipo aleatorio, o de acuerdo con un sistema de prioridad previamente establecido.

El mecanismo de servicio consiste en una o más instalaciones de servicio, con cada una de ellas con uno o más canales de servicios, llamados *servidores*. Los clientes son atendidos en estos servidores. El tiempo que transcurre desde el inicio del servicio para un cliente hasta su terminación se llama el *tiempo de servicio* (o duración del servicio). Un modelo de sistema de colas tiene que especificar la distribución de probabilidad de los tiempos de servicio de cada servidor (y tal vez para distintos tipos de clientes), aunque normalmente se supone la misma distribución para todos los servidores. Una vez más, la distribución exponencial es la más empleada en los tiempos de servicio.

6.2. Objetivos de la gestión de colas

En los modelos de colas existen dos objetivos: por un lado la minimización del tiempo de espera y por el otro la minimización de los costes totales de funcionamiento del sistema. Estos objetivos suelen ser conflictivos, ya que para reducir el tiempo de espera se necesitan poner más recursos en el sistema, con el consiguiente aumento de los costes de producción. En muchos casos el tiempo de espera es difícil de determinar, sobre todo cuando se trata de un sistema en donde seres humanos están implicados. En la figura 6.3 podemos ver la disyuntiva entre el coste de espera y el coste de producción.

FIGURA 6.3
Costes de un sistema de colas

Si pudiéramos sumar ambos costes, el coste total alcanzaría su mínimo en el punto H. En este punto el nivel de servicio es óptimo. Sin embargo, en muchos casos la obtención «objetiva» de este resultado puede ser muy complicada, ya que, como se ha indicado anteriormente, la cuantificación del tiempo de espera en valores monetarios puede ser harto complicada y subjetiva.

Por lo tanto, en general se intenta llegar a una solución que sea «lógica» en función de los valores que adopten los diferentes parámetros del modelo. En la sección siguiente se examinan estos parámetros.

6.3. Medidas del sistema

Existen dos tipos de medidas para poder valorar un sistema en donde pueden aparecer colas: medidas «duras» y medidas «blandas». Estas últimas están relacionadas con la calidad del servicio. Por ejemplo, no es lo mismo esperar quince minutos de pie haciendo cola en un ambulatorio sin refrigeración y poco ventilado que esperar el mismo tiempo en una sala de espera con butacas confortables, revistas, aire acondicionado y música clásica de fondo. El paciente valorará mucho más un minuto de espera en el primer caso, ya que representa un coste mucho más elevado en términos de confort. En otras palabras, seguramente un minuto de cola en el ambulatorio equivale a muchos minutos de espera en la sala de espera confortable. La gestión cuantitativa de las colas no se ocupa de estos aspectos cualitativos (que no por ello dejan de ser importantes), sino que da valores a una serie de medidas «frías» o «duras». Las medidas duras más utilizadas en los modelos de gestión de colas y su notación estándar son las siguientes:

- Número medio de personas en la cola, λ
- Número medio de personas en el sistema, μ
- Tiempo medio de espera en la cola, W_q
- Tiempo medio de estancia en el sistema, W_s
- Número medio de personas en la cola, L_q
- Número medio de personas en el sistema, W_s
- Porcentaje de ocupación de los servidores, P_w
- Probabilidad de que hayan x personas en el sistema, P_x

En los siguientes apartados iremos examinando estos conceptos.

6.4. Un sistema de colas elemental: tasa de llegada y de servicio constantes

Supongamos que tenemos un sistema en donde tanto la tasa de llegada (en personas por unidad de tiempo) como el tiempo de servicio son constantes. En este caso, podemos tener las tres situaciones siguientes:

6.4.1. *No hay cola, tiempo ocioso del servidor*

Supongamos que tenemos un sistema en donde cada seis minutos, exactamente, llega una persona a un ambulatorio. O, en otras palabras, la tasa de llegada es exactamente de 10 personas por hora. Supongamos que la tasa de servicio del médico (del servidor en términos técnicos) es de 12 personas por hora siempre, ni una más ni una menos. En esta situación nunca se formará una cola porque el servidor puede manejar perfectamente las llegadas. Incluso ya sabemos que el servidor estará ocioso un 16,6 % de su tiempo, ya que las llegadas necesitan únicamente de 10/12, o 83,33 % de la capacidad de servicio.

6.4.2. *No hay cola ni tiempo ocioso del servidor*

Siguiendo el ejemplo anterior, supongamos que la tasa de servicio pasa a ser igual a 10 personas por hora, es decir, exactamente igual qué la tasa de llegada. En esta situación es imposible que se forme una cola, pero por otro lado el servidor estará ocupado 100 % de su tiempo y trabajará a plena capacidad.

6.4.3. *Formación de cola y sin tiempo ocioso en el servidor*

Ahora supongamos que la tasa de servicio pasa a ser igual a ocho personas por hora, mientras que siguen llegando pacientes cada seis minutos exactamente. En esta situación se formará una cola que irá creciendo, ya que el servidor no puede absorber toda la demanda de servicio y los pacientes se irán acumulando. La cola de llegadas no servidas inmediatamente irá creciendo a una tasa de 2 personas por hora, es decir, el exceso de llegadas partido por las personas servidas. Por ejemplo, al cabo de ocho horas, tendríamos 16 personas en la cola.

El hecho de que hayamos asumido unas tasas de llegada y de servicio constantes hasta ahora facilita los cálculos para obtener información sobre el sistema. Pero la situación se complica si nos trasladamos a la situación más realista, en donde las tasas de llegada y de servicio no son constantes, sino que siguen una determinada distribución probabilística. Por ejemplo, si las llegadas y los tiempos de servicio estuviesen distribuidos aleatoriamente a lo largo de la jornada, aunque la capacidad de los servidores sea suficiente para absorber la demanda, puede pasar que un grupo de pacientes llegue en bloque y formen durante un tiempo una cola. Y, por otro lado, si durante un tiempo no llegan más pacientes, la cola puede ser reducida por el mecanismo de servicio. En las siguientes secciones examinaremos algunos de estos casos.

6.5. Las distribuciones de Poisson y exponencial

6.5.1. *La distribución de Poisson*

Esta distribución es muy frecuente en los problemas relacionados con la investigación operativa, sobre todo en el área de la gestión de colas. Suele describir, por ejemplo, la llegada de pacientes a un ambulatorio, las llamadas a una centralita telefónica, la llegada de coches a un túnel de lavado, el número de accidentes en un cruce, etc. Todos estos ejemplos tienen un punto en común: todos ellos pueden ser descritos por una variable aleatoria discreta que tiene valores no-negativos enteros (0,1,2,3,4...). El número de pacientes que llegan al ambulatorio en un intervalo de quince minutos puede ser igual a 0, 1, 2, 3...

Sigamos con el ejemplo del ambulatorio. La llegada de pacientes se puede caracterizar de la forma siguiente:

1. El número medio de llegadas de los pacientes para cada intervalo de quince minutos puede ser obtenido a través de datos históricos.

2. Si dividimos el intervalo de quince minutos en intervalos mucho más pequeños (por ejemplo, un segundo), podemos afirmar que:

   2.1. La probabilidad de que exactamente un único paciente llegue al ambulatorio por segundo tiene un valor muy reducido y es constante para cada intervalo de un segundo.

   2.2. La probabilidad de que dos o más pacientes lleguen dentro del intervalo de un segundo es tan pequeña que podemos decir que es igual a 0.

   2.3. El número de pacientes que llegan durante el intervalo de un segundo es independiente de dónde se sitúa este intervalo dentro del período de quince minutos.

   2.4. El número de pacientes que llegan en un intervalo de un segundo no depende de las llegadas que han sucedido en otro intervalo de un segundo.

Si al analizar un proceso de llegada éste cumple estas condiciones, podemos afirmar que su distribución es de Poisson.

La fórmula para obtener la probabilidad de que un evento ocurra (que lleguen tres pacientes, por ejemplo) es la siguiente:

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

En donde x representa el número de llegadas, λ la tasa media de llegadas y P(x) la probabilidad de que el número de llegadas sea igual a x.

6.5.2. *La distribución exponencial*

Mientras que la distribución de Poisson describe las llegadas por unidad de tiempo, la distribución exponencial estudia el tiempo entre cada una de estas llegadas. Si las llegadas son de Poisson, el tiempo entre ellas es exponencial. Mientras que la distribución de Poisson es discreta, la distribución exponencial es continua, porque el tiempo entre llegadas no tiene por qué ser un número entero.

Esta distribución se utiliza mucho para describir el tiempo entre eventos, más específicamente, la variable aleatoria que representa el tiempo necesario para servir a la llegada. Ejemplos típicos de esta situación son el tiempo que un médico dedica a una exploración, el tiempo de servir una medicina en una farmacia, o el tiempo de atender a una urgencia.

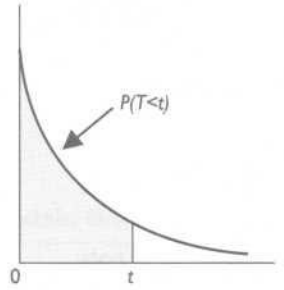
El uso de la distribución exponencial supone que los tiempos de servicio son aleatorios, es decir, que un tiempo de servicio determinado no depende de otro servicio realizado anteriormente, ni de la posible cola que pueda estar formándose. Otra característica de este tipo de distribuciones es que no tienen «edad», o en otras palabras, «memoria». Por ejemplo, supongamos que el tiempo de atención de un paciente en una sala quirúrgica sigue una distribución exponencial. Si el paciente ya lleva cinco horas siendo operado, la probabilidad de que esté una hora más es la misma que si hubiera estado dos horas, o diez horas o las que sea. Esto es debido a que la distribución exponencial supone que los tiempos de servicio tienen una gran variabilidad. A lo mejor el próximo paciente operado tarda una hora porque su cirugía era mucho más simple que la del anterior.

La función de densidad de la distribución exponencial es la siguiente:

$$p(t) = \mu e^{-ut}$$

En donde t representa el tiempo de servicio y μ la tasa media de servicio (pacientes servidos por unidad de tiempo). La densidad exponencial se presenta en la figura 6.4. En general nos interesará encontrar $P(T \leq t)$, la probabilidad de que el tiempo de servicio T sea inferior o igual a un valor específico t. Este valor es igual al área por debajo de la función de densidad.

FIGURA 6.4
Distribución exponencial



Si, por ejemplo, queremos saber cuál es la probabilidad de que el tiempo de servicio sea de dos o menos horas cuando el tiempo medio es de tres horas (una tasa de servicio de 1/3), podemos aplicar la fórmula siguiente:

$$P(T \leq t) = 1 - e^{-ut}$$

En este caso, $P(T \leq 2) = 0{,}486$, casi un 50 % de probabilidad.

6.6. Modelo de colas simple: llegadas en Poisson y tiempos de servicio exponencialmente distribuidos

El modelo que presentaremos a continuación tiene que cumplir las condiciones siguientes:

1. El número de llegadas por unidad de tiempo sigue una distribución de Poisson.
2. Los tiempos de servicio siguen una distribución exponencial.
3. La disciplina de la cola es de tipo FIFO.
4. La población potencial es infinita.

5. Existe un único canal de servicio.
6. La tasa media de llegadas es menor que la tasa media del servicio.
7. El tamaño potencial de la cola es infinito.

Si estas condiciones se cumplen y si conocemos la tasa media de llegada λ, y la tasa media de servicio μ, las ecuaciones para obtener valores de las medidas descritas anteriormente son:

- Número medio en la cola:

$$L_q = \frac{\lambda^2}{\mu \mu - \lambda}$$

- Número medio en el sistema:

$$L_s = \frac{\lambda}{\mu - \lambda}$$

- Tiempo medio de espera en la cola:

$$W_q = \frac{\lambda}{\mu \mu - \lambda}$$

- Tiempo medio en el sistema:

$$W_s = \frac{1}{\mu - \lambda}$$

- Factor de utilización:

$$P_w = \frac{\lambda}{\mu}$$

Un ejemplo

Consideremos el caso de un gran laboratorio farmacéutico que tiene en su almacén un único estacionamiento de carga que sirve a todas las farmacias de una región, y existe un único trabajador para buscar los medicamentos del pedido de cada furgoneta y cargarlos en ella. Se observa que de vez en cuando las furgonetas de transporte se acumulan en el estacionamiento formando cola, y de vez en cuando el trabajador está ocioso. Después de un estudio del sistema observamos que éste cumple las condiciones expuestas anteriormente. Después de examinar las llegadas de las camionetas durante varias semanas, se determina que la tasa media de llegada es de cuatro camionetas por hora, y que la tasa de servicio es de 6 camionetas por hora. Los gestores del almacén están considerando el añadir un trabajador adicional, o incluso dos de ellos, para aumentar la tasa de servicio. El problema consiste en evaluar estas opciones diferentes.

Si se añade un trabajador, el sistema seguirá siendo de cola simple, porque una única camioneta puede cargarse a la vez. Si usamos dos trabajadores, la tasa de servicio será igual a 12. Si utilizamos tres trabajadores, la tasa de servicio será igual a 18.

En el cuadro 6.1 se han utilizado las ecuaciones expuestas anteriormente para obtener las medidas de eficiencia del sistema. Hemos supuesto que la capacidad de trabajo es proporcional al número de trabajadores.

CUADRO 6.1
Resultados del modelo simple de colas

		Trabajadores		
		1	2	3
Número medio de camionetas en la cola	L_q	1,333	0,167	0,063
Número medio de camionetas en el sistema	L_s	2,000	0,500	0,286
Tiempo medio de la camioneta en cola	W_q	0,333	0,042	0,016
Tiempo medio de la camioneta en el sistema	W_s	0,500	0,125	0,071
Ocupación del servicio	P_w	0,667	0,333	0,222

Supongamos que los costes de operación de cada camioneta por hora son de 2.000 ptas y los trabajadores cobran 1.800 ptas por hora de trabajo y que éstos trabajan 8 horas al día. En el cuadro 6.2 se presentan los costes asociados. Al interpretar los tiempos, hay que ir con cuidado, ya que éstos están en fracciones de hora.

CUADRO 6.2
Costes de operación del sistema

Trabajadores	Coste de camioneta por día	Coste de mano de obra por día	Coste total por día
1	320.000	144.000	464.000
2	80.000	288.000	368.000
3	46.000	432.000	478.000

Los gestores tendrían que añadir un nuevo trabajador al sistema, ya que esto representará una reducción de los costes totales operacionales, aunque el factor de utilización pasará a ser de 33 %. Es decir, que los dos trabajadores tendrán 5 horas y 20 minutos para dedicarse a otras tareas dentro del laboratorio farmacéutico.

Extensión del modelo simple a colas con capacidad limitada
Existen casos en los que el sistema (cola más servicio) tiene una cierta capacidad. Si un cliente llega cuando hay M o más personas en el sistema, el cliente se va inmediatamente y no vuelve. Este tipo de modelo es característico de los problemas de colas que se pueden encontrar en algunos servicios. Por ejemplo, un restaurante con un estacionamiento limitado. En este caso, las ecuaciones del modelo son:

- Probabilidad de 0 personas en el sistema:

$$P_0 = \frac{1-(\lambda/\mu)}{1-\lambda/\mu^{M+1}}$$

- Factor de utilización:

$$P_w = 1 - P_0$$

- Proporción de clientes perdidos porque el sistema está lleno:

$$P_M = \lambda/\mu^{\;M} P_0$$

- Número medio en el sistema:

$$L_s = \frac{P_w - M\lambda/\mu P_M}{1-\lambda/\mu}$$

- Número medio en la cola:

$$L_q = L_s - \frac{\lambda 1 - P_M}{\mu}$$

- Tiempo medio de espera en el sistema:

$$W_s = \frac{L_s}{\lambda 1 - P_M}$$

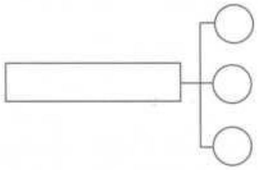
- Tiempo medio en la cola:

$$W_q = W_s - \frac{1}{\mu}$$

6.7. Modelo múltiple de colas: llegadas en Poisson y tiempos de servicio exponencialmente distribuidos

En muchos casos podemos tener situaciones en donde existe más de un servidor en el sistema. A medida que van llegando los clientes, los servidores se van ocupando y cada vez que uno de ellos acaba su servicio, el primero de la cola lo vuelve a ocupar. El sistema está representado en la figura 6.5.

FIGURA 6.6
Sistema múltiple de colas



En este tipo de modelos la tasa de llegada siempre tiene que ser inferior a la tasa agregada de servicio, que no es más que la tasa de servicio individual multiplicada por el número de canales. En este modelo se supone, además de las condiciones expuestas anteriormente, que la tasa individual de cada canal es la misma. Las expresiones matemáticas para la obtención de las medidas de eficiencia del sistema dependen de P_0, que es la probabilidad de que no haya nadie en el sistema (cola más servicio). En el disquete (archivo colas.xls) se pueden calcular los valores de P_0. Los valores de las medidas de eficiencia son función de P_0 y se obtienen a partir de las siguientes fórmulas:

- Factor de utilización:

$$P_W = \frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k \frac{k\mu}{k\mu - \lambda} P_0$$

- Número medio en el sistema:

$$L_S = \frac{\lambda\mu \; \lambda/\mu^{k}}{k-1! \; k\mu-\lambda^{2}} P_0 + \frac{\lambda}{\mu}$$

- Número medio en la cola:

$$L_q = L_S - \frac{\lambda}{\mu}$$

- Tiempo medio de espera en el sistema:

$$W_S = \frac{L_S}{\lambda}$$

- Tiempo medio en la cola:

$$W_q = \frac{L_q}{\lambda}$$

Tenemos que tener presente que en estas ecuaciones μ representa la tasa de servicio por canal. En el disquete adjunto (archivo colas.xls) se pueden calcular estos valores tanto para sistemas múltiples de colas con capacidad infinita como con capacidad limitada.

Un ejemplo

El ambulatorio de una región tiene dos médicos de cabecera que atienden a los pacientes que van llegando. En general, los pacientes tienen que esperar a ser atendidos y la gerencia está estudiando la posibilidad de contratar un nuevo médico para aligerar el sistema. Como es muy difícil estimar en términos monetarios el coste de espera de los pacientes, la gerencia realizará la nueva contratación si se consigue reducir los tiempos totales del servicio (espera más atención) a la mitad. Después de observar y recoger datos sobre las llegadas y sobre el tiempo de servicio, la gerencia calcula que en media llegan ocho pacientes por hora, y que cada uno de los médicos puede atender cinco pacientes por hora.

En el cuadro 6.3 se presentan los resultados después de aplicar las fórmulas del modelo con dos y tres médicos.

CUADRO 6.3
Resultados del modelo múltiple de colas

		Médicos	
		2	3
Probabilidad de que todos los médicos estén libres	P_0	0,111	0,190
Probabilidad de que todos los médicos estén ocupados	P_w	0,710	0,278
Número medio de pacientes en el sistema	L_s	4,442	1,918
Número medio de pacientes en cola	L_q	2,842	0,318
Tiempo medio de un paciente en el sistema	W_s	0,555	0,240
Tiempo medio de un paciente en cola	W_q	0,355	0,040

En el cuadro podemos observar que, si añadimos un médico adicional, el tiempo de espera de cada paciente en el sistema pasa de 0,555 horas a 0,240 horas. Por lo tanto, el objetivo de la gerencia se cumple al añadir un nuevo médico. También se puede observar que con tres médicos el tiempo de espera en la cola es insignificante.

6.8. Limitaciones de los modelos de gestión de colas

Los dos modelos que hemos presentado en este capítulo son los más comunes cuando se trata de sistemas en donde están implicados seres humanos. Sin embargo, pueden existir casos en donde la población potencial del sistema es finita, la cola de la disciplina no es FIFO, la tasa de servicio depende de las personas en la cola, y las distribuciones de las llegadas no son de Poisson. En estos casos estos modelos son inservibles.

Las distribuciones juegan un papel esencial en estos modelos. Los sistemas en donde las variaciones de las llegadas en diferentes horarios son muy grandes no pueden ser examinados con las formulaciones presentadas. Cuando tenemos sistemas más complejos se utiliza la simulación como método de análisis. En la siguiente sección simularemos un sistema de colas.

6.9. Ejemplo de simulación de un sistema de colas

La farmacia de un hospital tiene dos personas para atender a 10 enfermeras que vienen a buscar medicamentos para los pacientes. La gerencia observa que de vez en cuando se forman colas para recoger las medicinas y que el servicio resulta un tanto ineficiente. Por otro lado, la población potencial es bastante reducida y no puede considerarse como infinita. Por otro lado, la distribución del número de llegadas no es Poisson ni el tiempo de servicio exponencial. Por lo tanto, no se puede aplicar el modelo múltiple de gestión de colas.

6.9.1. *Recogida de datos*

La gerencia observó el funcionamiento de la farmacia durante períodos de una hora distribuidos a lo largo de un mes. Estos períodos de una hora fueron aleatoriamente escogidos durante el día para obtener una muestra representativa de la actividad.

CUADRO 6.4
Resultados de la muestra

Duración del tiempo de servicio (minutos)	Número de Observaciones
8	15
9	30
10	45
11	60
Total de llegadas	150

Los resultados de la observación se muestran en el cuadro 6.4.

Además, la gerencia dividió el tiempo de observación en intervalos de cinco minutos y anotó las llegadas de enfermeras que llegaron durante estos intervalos. Se observó que en promedio llegaba una enfermera cada cinco minutos. Al final del período de observación, la gerencia presentó los resultados obtenidos de la siguiente forma:

- Distribución porcentual de los tiempos de servicio:

$$15/150 = 10\ \%\ (8\ \text{min})$$
$$30/150 = 20\ \%\ (9\ \text{min})$$
$$45/150 = 30\ \%\ (10\ \text{min})$$
$$60/150 = 40\ \%\ (11\ \text{min})$$

- Media ponderada de los tiempos de servicio:

10 % × 8 min =	0,8 min
20 % × 9 min =	1,8 min
30 % × 10 min =	3,0 min
40 % × 11 min =	4,4 min
Tiempo medio de servicio:	10,0 min

Con esta información, la gerencia pudo empezar a realizar la simulación con la ayuda de números aleatorios (ver anexo 2).

6.9.2. *Simulación de llegadas*

En primer lugar la gerencia simula las llegadas de las enfermeras a la farmacia. Ésta sabe que las llegadas son aleatorias, aunque en promedio llega una cada cinco minutos. Como los números aleatorios tienen 10 dígitos (0,1,2,3,4,5,6,7,8,9), la gerencia escoge (aleatoriamente) el 7 como representativo de una llegada. Si cogemos al azar un número de la tabla, la cantidad de sietes que contenga el número indicará la cantidad de llegadas en un intervalo de 5 minutos. La gerencia simula las llegadas a la farmacia durante 24 períodos de cinco minutos. Quizás no sea una cantidad muy representativa en este caso, pero para efectos de explicación de la simulación es suficiente. En un caso real simularíamos el sistema con muchos más períodos, pero la mecánica seguiría siendo la misma.

Para ilustrar el procedimiento de simulación de llegadas, hemos escogido los 12 primeros números aleatorios del apéndice 2 (por columna) y contado las veces que sale el número 7.

1239650125	0	6749281769	2	0178780337	3
1370937859	2	8912349495	0	9128374452	1
0926561938	0	9172674928	2	4412773934	2
1639438732	1	9916253764	0	0112378549	1

En el cuadro 6.5 se muestran los resultados para los 24 períodos.

CUADRO 6.5
Simulación del número de llegadas en cada intervalo

Período	Cantidad de llegadas	Período	Cantidad de llegadas
1	0	13	0
2	2	14	0
3	0	15	1
4	1	16	4
5	2	17	1
6	0	18	1
7	2	19	1
8	0	20	0
9	3	21	0
10	1	22	1
11	2	23	0
12	1	24	2

6.9.3. *Simulación de los tiempos de servicio*

Después de obtener los resultados de la simulación de las llegadas, la gerencia tiene que realizar la simulación de los tiempos de servicio. Recordemos la distribución de los tiempos de servicios observada anteriormente:

Minutos	Porcentaje
8	10
9	20
10	30
11	40

Como seguimos utilizando los números aleatorios de 10 dígitos, consideraremos que el 0 representa un tiempo de servicio de 8 minutos, el 1 y el 2 un tiempo de 9 minutos, el 3, 4 y el 5 un tiempo de 10 minutos, y el 6, 7, 8 y 9 un tiempo de 11 minutos. De esta forma podemos representar exactamente la probabilidad de los tiempos de llegada. Por ejemplo, observamos en el cuadro 7.5 que en el segundo período hubieron dos llegadas al servicio. Para simular el tiempo de servicio, escogemos la última fila de números aleatorios de la tabla comenzando por la izquierda. El primer número es 9 y el segundo 8. Esto quiere decir que la primera y la segunda llegada tendrán asociadas un tiempo de atención de 11 minutos. Este proceso se repite para todos los períodos en los que hay llegadas. El resultado se presenta en el cuadro 6.6.

6.9.4. *Simulación conjunta del sistema*

Ahora ya podemos simular el sistema. El objetivo de la gerencia es encontrar el número óptimo de trabajadores en la farmacia de forma a minimizar el coste total del servicio. Se considera que el sistema es FIFO, es decir, primero en llegar, primero en ser atendido. Ahora se tienen que definir los criterios de cada llegada dentro de cada período. Se definen los criterios siguientes:

1. Si en un período llega una única enfermera, ésta lo hará al principio del período.
2. Si en un período llegan 2 enfermeras, la primera lo hace al principio y la segunda en el tercer minuto.
3. Si en un período llegan 3 enfermeras, la primera llega al principio, la segunda en el minuto 3 y la tercera en el minuto 5.
4. Si en un período llegan 4 enfermeras, se asume que llegan en los minutos 2, 3, 4 y 5, respectivamente.

CUADRO 6.6
Resultados de la simulación de los tiempos de servicio

Número de período	Número de Llegadas	Tiempo de servicio de cada una
1	0	
2	2	(1) 11 min, (2) 11 min
3	0	
4	1	(3) 10 min
5	2	(4) 11 min, (5) 10 min
6	0	
7	2	(6) 9 min, (7) 10 min
8	0	
9	3	(8) 10 min, (9) 10 min, (10) 11 min
10	1	(11) 11 min
11	2	(12) 10 min, (13) 8 min
12	1	(14) 11 min
13	0	
14	0	
15	1	(15) 11 min
16	4	(16) 10 min, (17) 11 min, (18) 11 min, (19) 11 min
17	1	(20) 9 min
18	1	(21) 8 min
19	1	(22) 11 min
20	0	
21	0	
22	1	(23) 11 min
23	0	
24	2	(24) 9 min, (25) 11 min

En la figura 7.4 se presenta el patrón de llegadas de los 24 períodos.

Cada llegada viene indicada por un cuadro conteniendo su correspondiente número de orden, y encima del recuadro se indica el tiempo de espera correspondiente. Por ejemplo, si examinamos el período entre las 10:15 y las 10:20, vemos que se producen 5 llegadas (de la 16 a la 20), y, como veremos más adelante, seguramente se producirá una cola considerable. El comportamiento del sistema simulado con dos trabajadores atendiendo a los clientes (nivel de congestión, personas en cola, duración del tiempo de espera) puede representarse tal como se muestra en la figura 7.5. En la figura 3.6 se representa la simulación con tres trabajadores. Si se comparan las dos figuras, visualmente se puede observar que el tiempo de espera se reduce considerablemente.

Si se examina desde un punto de vista económico, con dos trabajadores atendiendo a las enfermeras, éstas esperarán un total de 213 minutos, un tiempo medio de espera igual a 8,52 minutos. Para obtener un valor monetario del tiempo de espera, la gerencia considera que el coste por hora de cada trabajador es igual a 7 euros y el coste de cada enfermera es de 12 euros. Si recordamos que el tiempo medio entre cada llegada era de 5 minutos, en media las enfermeras realizarán 96 viajes por día (8 horas diarias por 12 viajes por hora). Y si el tiempo medio de espera es de 8,52 minutos por viaje, el tiempo total de espera es igual a 817,9 minutos (13,63 horas perdidas), lo que representa un coste total de espera de 163,56 euros. Si añadimos el coste de los dos trabajadores (112 euros), el coste diario total es igual a 275,56 euros.

Si realizamos el mismo ejercicio, pero modificando el número de trabajadores atendiendo a las enfermeras (figura 7.6), el tiempo total de espera es igual a 47 minutos, o 1,88 minutos de espera por llegada. Si tenemos 96 llegadas por día, el tiempo total perdido es igual a 180,48 minutos (3 horas diarias). El coste de la espera es igual a 36 euros y el sueldo de los trabajadores igual a 168 euros, lo que da un coste total igual a 204 euros. Por lo tanto, la operación con tres trabajadores parece ser más eficiente en términos monetarios. ¿Qué pasaría si contratáramos un cuarto trabajador que eliminaría completamente el tiempo de espera de las enfermeras? En este caso el único coste sería el sueldo de los trabajadores, que sería igual a 224 euros, superior al coste con tres trabajadores.

FIGURA 6.6
Representación de las llegadas

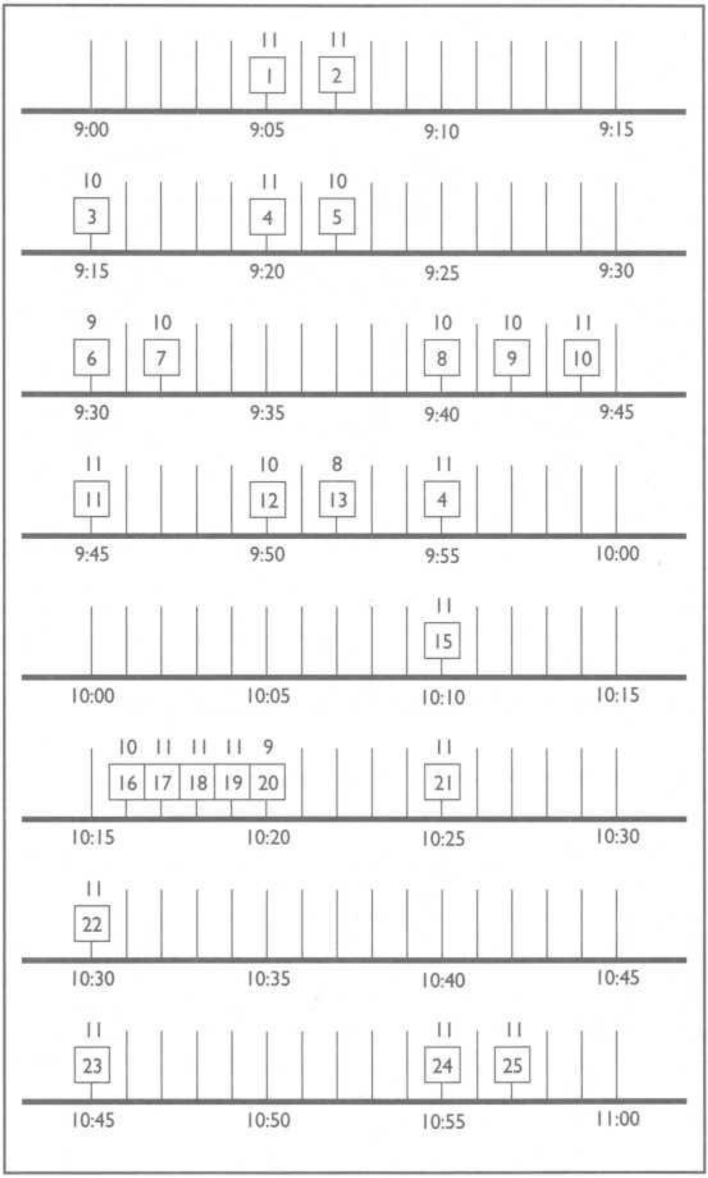


FIGURA 6.7
Operación de la farmacia con dos trabajadores

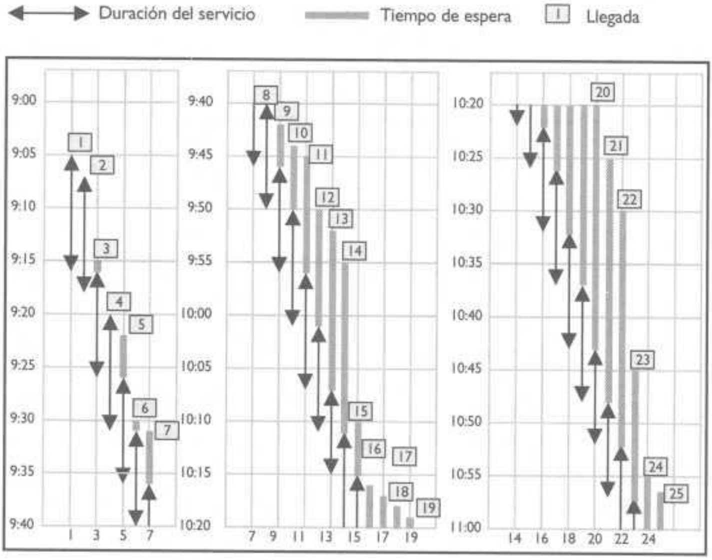
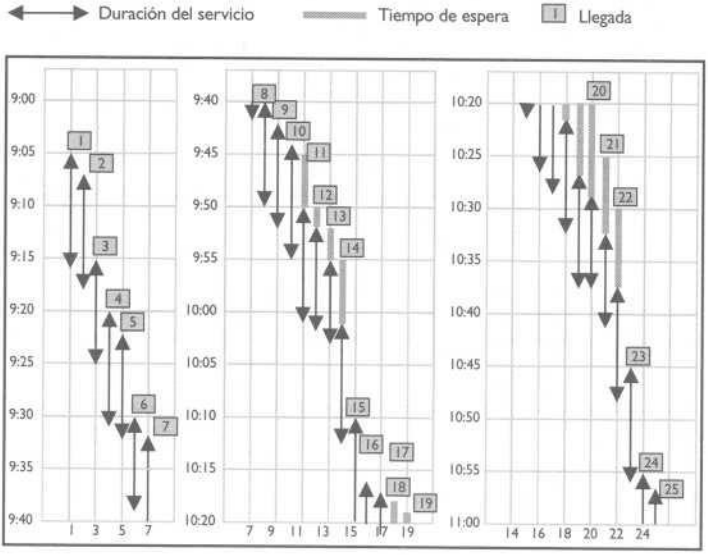


FIGURA 6.8
Operación de la farmacia con dos trabajadores



6.10. Problemas

6.1. En un ambulatorio se reciben una media de tres pacientes por hora, siguiendo una distribución de Poisson. El único médico que está en el ambulatorio atiende una media de 6 pacientes por hora. Los tiempos de atención siguen una distribución exponencial. La pregunta que se plantea el gestor es si vale la pena contratar a un nuevo médico o no (considerando que éste realizará el mismo número de pacientes).

6.2. La llegada de enfermeras a una farmacia del hospital Todosalud puede describirse a través de una distribución de Poisson. Los tiempos de servicio utilizan una distribución exponencial. La tasa de llegada es de 45 enfermeras por hora, mientras que cada farmacéutico puede atender a 50 enfermeras por hora. El coste de cada enfermera es de 15 euros por hora, mientras que cada farmacéutico gana 10 euros por hora. Encontrar el número óptimo de farmacéuticos a contratar.

6.3. Un centro de atención primaria tiene que administrar la vacuna de poliomielitis a los niños de un barrio. El centro está organizado de forma que los padres van llegando con los niños, formando una cola y siendo atendidos 40 por hora, con una distribución exponencial, por cualquiera de las enfermeras que están de servicio. Este servicio de vacunación se ofrece una vez a la semana, y en este día las llegadas se realizan con una tasa igual a 40 niños por hora. El director del centro sabe que la mayoría de los padres vienen durante sus horas de trabajo y por ello quiere limitar el tiempo total de administración de la vacuna a 15 minutos (incluyendo la espera). ¿Cuántas enfermeras tendrá que utilizar el gerente?

6.4. Un fisioterapeuta tiene un pequeño consultorio en donde trabaja con pacientes operados de rodilla. Son en media 8 los pacientes que llegan por las tardes siguiendo una distribución de Poisson. El tiempo de servicio del fisioterapeuta sigue una distribución exponencial con una tasa media de servicio de 9 pacientes por hora. Por las mañanas no está nada ocupado, porque la mayoría de pacientes prefiere venir por las tardes. Después de varios años de experiencia, el fisioterapeuta sabe que, en general, si un paciente llega a su clínica y hay otros tres esperando a ser atendidos, éste cambiará de fisioterapeuta. Para intentar convencer a algunos de sus pacientes a venir por la mañana, decide dar vales para almorzar en el bar de la esquina. Con esta táctica cree que podrá reducir la tasa de llegadas de la tarde a 5 pacientes por hora. Si la pérdida de un paciente significa una pérdida de 20 euros, ¿cuánto estaría dispuesto a pagar por cada vale considerando que cada mañana y cada tarde es de 4 horas?

APÉNDICE 6.1
Tabla de números aleatorios

1239650125	1142160902	4527058366	1430875177	2821660618	8377041005
1370937859	8611794741	9815467859	9902648765	8411668944	7446544112
0926561938	1497952703	8609672743	4339955985	7272012830	1644676720
1639438732	2405306350	7914214780	4677121705	7791648017	8176854669
6749281769	7502962776	9671479586	9093449703	5368210111	4339625930
8912349495	7679203454	9981241957	4032184023	4683141252	2008225353
9172674928	7759918425	7230206493	2581236968	3177342357	9568401934
9916253764	2363211025	6282095638	3001769692	8059067287	4185944112
0178780337	8652658498	7424574963	5185675861	9919589016	7193770252
9128374452	9615307086	3542787227	6735835540	6005786976	1738020959
4412773934	5251827422	3399383961	9764534919	7108528926	2628877567
0112378549	2562566485	6582696326	1089265935	7552101346	6433830583
6484118917	4127133914	3687936583	8363383157	9380150844	2984673476
7485396877	1398185872	1298594465	5141064723	5658307618	3659683979
4449420399	3106053098	5647103564	4730042272	9397192556	6329548442
4931720558	5644540248	8583519254	5956489316	6647190158	6024106022
7019256915	8735782233	9315124105	6211324018	2835893430	3460537007
7088642561	3802788585	4015059652	3577040333	5596749322	6100976968
1322899809	8411711735	3383210443	1558138811	8636254576	9483732238
7644596455	7727801919	2425499018	9548164500	3705900236	1713960284
4350250689	5583998402	8094153554	7641705754	1338635552	7597030515
3969625810	7127183350	4809003690	2376632426	7457842856	6287196405
9867466169	3959776552	8642913942	3395147232	1594796171	6959239675
2336529609	8083742448	1499051196	8052499450	8186045265	1772149893
8497899420	6831646060	7649824834	3614066579	5186204268	3416788948
1383109138	8248598021	5880393712	1869970313	3950714593	7974854777
9755836769	8902846338	9965787689	7335817518	7384119119	3462832927
9459378972	9605697733	1001598997	2998482807	2873138839	7637508522
1156270803	1745774801	1544226029	7718016727	7240282667	6018667991
4281042330	4927846359	3928115864	5928523182	1770087219	1465417792

7. SOLUCIONES A LOS PROBLEMAS

7.1. Capítulo 2

Problema 2.1

- P_1^A = Cantidad de pacientes de tipo 1 tratados en el centro A.
- P_1^B = Cantidad de pacientes de tipo 1 tratados en el centro B.
- P_2^B = Cantidad de pacientes de tipo 2 tratados en el centro B.
- P_1^C = Cantidad de pacientes de tipo 1 tratados en el centro C.
- P_2^C = Cantidad de pacientes de tipo 2 tratados en el centro C.

Formulación del problema:

$$
\begin{array}{lllllll}
\text{Min } Z = & 26P_1^A & +28P_1^B & +33P_2^B & +24P_1^C & +28P_2^C & \\
\text{s. a.} & & & & & & \\
& P_1^A & +P_1^B & & +P_1^C & & = 100 \\
& & +P_2^B & & & +P_2^C & = 150 \\
& P_1^A & & & & & \leq 80 \\
& & +P_1^B & +P_2^B & & & \leq 50 \\
& & & & +P_1^C & +P_2^C & \leq 120 \\
\end{array}
$$

$$P_1^A, P_1^B, P_2^B, P_1^C, P_2^C \geq 0$$

Problema 2.2

- T = número de anuncios en televisión.
- R = número de anuncios en radio.

Formulación del problema:

$$
\begin{array}{lrrl}
\text{Max } Z = & 25T & +R & \\
 & 100T & +5R & \leq 1.000 \\
 & 2T & -R & \leq 0 \\
 & & T, R \geq 0 &
\end{array}
$$

Problema 2.3

- P = pacientes de pediatría.
- D = pacientes de dermatología.
- O = pacientes de ortopedía.
- N = pacientes de neurocirugía.

$$
\begin{array}{lrrrrll}
\text{Max } Z = & 100P & +200D & +150O & +180N & & \\
\text{s. a.} & & & & & & \\
 & P & & +O & & \leq & 120 \\
 & 5P & +5D & +2O & +4N & \leq & 200 \\
 & 2P & +8D & +O & +5N & \leq & 140 \\
 & P & +D & & +8N & \leq & 110 \\
 & 4P & +8D & +16O & +10N & \leq & 240 \\
 & 10P & +14D & +8O & +12N & \leq & 320 \\
 & & & P, D, O, N \geq 0 & & &
\end{array}
$$

Problema 2.4

- A_R = Cantidad, en mg, de ácido acetil-salicílico en Resacón.
- A_J = Cantidad, en mg., de ácido acetil-salicílico en Jaquecón.
- P_R = Cantidad, en mg, de paracetamol en Resacón.
- P_J = Cantidad, en mg, de paracetamol en Jaquecón.

$$
\begin{array}{lrrrrl}
\text{Max } Z = & 70A_R & +120A_J & +30_{PR} & +80P_J & \\
\text{s. a.} & & & & & \\
 & -0{,}25A_R & & +0{,}25_{PR} & & \leq 0 \\
 & & -0{,}75A_J & & +0{,}25P_J & \leq 0 \\
 & & 0{,}50A_J & & -0{,}50P_J & \leq 0 \\
 & A_R & & P_R & & \leq 400 \\
 & & A_J & & P_J & \leq 300 \\
 & & & A_R, A_J, P_R, P_J \geq 0 & &
\end{array}
$$

Problema 2.5

- X_A = pacientes de tipo X tratados en el hospital A.
- Y_A = pacientes de tipo Y tratados en el hospital A.
- Z_A = pacientes de tipo Z tratados en el hospital A.
- X_B = pacientes de tipo X tratados en el hospital B.
- Y_B = pacientes de tipo Y tratados en el hospital B.
- Z_B = pacientes de tipo Z tratados en el hospital B.
- X_C = pacientes de tipo X tratados en el hospital C.
- Y_C = pacientes de tipo Y tratados en el hospital C.
- Z_C = pacientes de tipo Z tratados en el hospital C.

$$
\begin{array}{lcccccccccccc}
\text{Max} = & 420X_A & +360Y_A & +300Z_A & +420X_B & +360Y_B & +300Z_B & +420X_C & +360Y\mathbf{C} & +300Z_C & & \\
\text{s.a.} & & & & & & & & & & & \\
& X_A & +Y_A & +Z_A & & & & & & & & \leq 700 \\
& & & & X_B & +Y_B & +Z_B & & & & & \leq 800 \\
& & & & & & & +X_C & +Y_C & +Z_C & & \leq 450 \\
& 2X_A & +1{,}5Y_A & +1{,}2Z_A & & & & & & & & \leq 13 \\
& & & & +2X_B & +1{,}5Y_B & +1{,}2Z_B & & & & & \leq 12 \\
& & & & & & & 2X_C & +1{,}5Y_C & +1{,}2Z_C & & \leq 5 \\
& 800X_A & +800Y_A & +800Z_A & -700X_B & -700Y_B & -700Z_B & & & & & \leq 0 \\
& X_A & Y_A & Z_A & X_B & Y_B & Z_B & X_C & Y_C & Z_C & 0 & \leq 0
\end{array}
$$

Problema 2.5

- X_i = cantidad invertida en el fondo de inversión i

$$\text{Max } Z = \sum_{i=1}^{20} (1 + r_i) X_i$$

s. a.

$$\sum_{i=1}^{20} X_i = 10.000.000$$

$$X_i \leq 2.000.000$$

$$X_i \geq 500.000$$

$$X_i \geq 0$$

Problema 2.6

Parámetros:

- i, m = índice y número de laboratorios.
- j, n = índice y número de almacenes.
- t, T = índice y número de trimestres.
- s_i = capacidad de *stock* máxima en el laboratorio i
- p_i = coste de producción unitario
- c_{ij} = coste unitario de transporte entre i y j

- l_i^t = capacidad de producción máxima del laboratorio i en el período t
- k_i^t = coste unitario de almacenaje en el laboratorio i
- h_j^t = coste unitario de almacenaje en el almacén j
- r_j^t = demanda que sale del almacén j durante el período t

Variables:

- X_i^t = cantidad producida en el laboratorio i y enviada en el período t
- S_i^t = cantidad producida en i y almacenada en el período t
- Z_{ij}^t = cantidad transportada del laboratorio i al almacén j en el período t
- W_j^t = cantidad almacenada en el almacén j al final del período t

Formulación:

$$\text{Min } Z = \sum_{t=1}^{T}\left[\sum_{i=1}^{m} k_i^t S_i^t + \sum_{j=1}^{n} h_j^t W_j^t + \sum_{i=1}^{m}\sum_{j=1}^{n} Z_{ij}^t + \sum_{i=1}^{m} X_i^t\right]$$

s. a.

$$\sum_{j=1}^{n} Z_{ij}^t - S_i^{t-1} - X_i^t \leq 0 \quad i=1,\dots,m \quad t=1,\dots,T$$

$$X_i^t + S_i^t \leq X_i^t \quad i=1,\dots,m \quad t=1,\dots,T$$

$$\sum_{i=1}^{m} Z_{ij}^t - W_{ji}^{t-1} \leq r_j^t \quad j=1,\dots,n \quad t=1,\dots,T$$

$$X_i^t, S_i^t, Z_{ij}^t, W_j^t \geq 0 \quad i=1,..,m \; j=1,\dots,n \; t=1,\dots,T$$

7.2. Capítulo 3

7.2.1. *Cierto - Falso*

1) C	6) C	11) C	16) F
2) F	7) C	12) C	17) F
3) F	8) F	13) C	18) C
4) C	9) F	14) F	19) C
5) F	10) C	15) F	20) C

7.2.2. *Elección Múltiple*

1) b	5) c	9) c	13) a
2) d	6) b	10) d	14) c
3) b	7) c	11) d	15) d
4) d	8) b	12) b	

7.2.3. *Problemas*

Problema 3.1

Puntos extremos

X_1	X_2	$Z = X_1 + 2X_2$	$Z = X_1 + 6X_2$
0	0	0	0
5	0	5	5
4	1	6	10
0	2	4	12

Problema 3.2

Definimos S_1 S_2 y S_3 como las variables de holgura asociadas las restricciones. 2 puntos extremos factibles:

$$X_1 = 0;\ X_2 = 0;\ X_3 = 0;\ S_1 = 12;\ S_2 = 11;\ S_3 = 20;\ Z = 0$$
$$X_1 = 3{,}09;\ X_2 = 2{,}23;\ X_3 = 0;\ S_1 = 0;\ S_2 = 0;\ S_3 = 7{,}13;\ Z = 14{,}59$$

Problema 3.3

Definimos S_1 S_2 y S_3 como las variables de holgura asociadas las restricciones. La solución óptima es: $X_1 = 6{,}67$; $X_2 = 0$; $X_3 = 36{,}67$; $S_1 = 116{,}6$; $S_2 = 0$; $S_3 = 0$; $Z = 66{,}67$

Problema 3.4

Definimos S_1 S_2 y S_3 como las variables de holgura asociadas las restricciones. Como sabemos que en la solución óptima X_1 y X_3 son positivas y $X_2 = 0$, podemos afirmar que, en la solución óptima, las dos primeras son básicas, mientras que la última no es básica. Al efectuar el método Simplex, en la iteración inicial las tres variables son no-básicas, por lo que escogeremos X_1 o X_3 para entrar en la base, pero nunca X_2. En cada iteración intentaremos que X_2 no entre en la base (porque al final tendrá que salir). La solución óptima del problema es: $X_1 = 10$; $X_2 = 0$; $X_3 = 10$; $S_1 = 0$; $S_2 = 4$; $S_3 = 0$; $Z = 50$.

Problema 3.5

Soluciones óptimas alternativas:

$$X_1 = 3;\ X_2 = 0;\ X_3 = 2;\ X_4 = 0;\ Z = 5$$
$$X_1 = 0;\ X_2 = 3;\ X_3 = 2;\ X_4 = 0;\ Z = 5$$
$$X_1 = 1{,}5;\ X_2 = 1{,}5;\ X_3 = 1;\ X_4 = 1;\ Z = 5$$

Problema 3.6

Solución óptima: $X_1 = 13$; $X_2 = 0$; $X_3 = 5$; $X_4 = 0$; $Z = 51$

Problema 3.7

Solución óptima:

$$P_1^A = 80; P_1^B = 0; P_2^B = 50; P_1^C = 20; P_2^C = 100; Z = 1010$$

Problema 3.8

Solución óptima: $T = 9{,}1$; $R = 18{,}2$; $Z = 245{,}5$

Problema 3.9

Solución óptima: $P = 100$; $D = 200$, $O = 150$, $N = 180$; $Z = 4608{,}6$

Problema 3.10

Solución óptima: $A_r = 300$; $A_l = 150$, $P_r = 100$, $P_l = 150$; $Z = 54000$

Problema 3.11

Solución óptima: $X_A = 3{,}8$; $Y_A = 1{,}3$; $Z_A = 0{,}8$; $X_B = 3{,}8$; $Y_B = 3{,}0$; $Z_B = 0$; $X_C = 0{,}6$; $Y_C = 0$; $Z_C = 6822$

7.3. Capítulo 5

Problema 5.1

Puntos eficientes encontrados.

w_1	w_2	x_1	x_2	z_1	z_2
1	0	0	2	2	–4
0,5	0,5	3	0	–6	9

Problema 5.2

Puntos eficientes diferentes encontrados con el método de los pesos:

w_1	w_2	x_1	x_2	x_3	z_1	z_2
1	0	3,1	2,2	0,0	7,5	5,8
0,7	0,3	0	3	0	6	12
0	1	0	2,1	3,8	0,3	19,6

Problema 5.3

Puntos eficientes diferentes encontrados con el método de los pesos:

w_1	w_2	x_1	x_2	x_3	z_1	z_2
1	0	6,7	0	36,7	66,7	–103,3
0,8	0,2	3,2	5,2	28	58	–54,8
0,7	0,3	0	5,6	24,4	54,4	–45,6
0,5	0,5	0	2,5	0	2	12,5

Problema 5.4

Puntos eficientes encontrados con el método de la restricción.

x_1	x_2	z_1	z_2
0	2	4	–4
0	1,5	3	–3
0	1	2	–2
0	0,5	1	–1
0	0	0	0
1	0	–3	1
2	0	–6	2

7.4. Capítulo 6

Problema 6.1

Tasa de llegada	3	3
Tasa de servicio	6	6
Número de servidores	1	2
Factor de utilización	0,500	0,100
Probabilidad de 0 en el sistema	0,500	0,600
Número medio en la cola	0,500	0,333
Tiempo medio en la cola	0,167	0,011
Número medio en el sistema	1,000	0,533
Tiempo medio en el sistema	0,333	0,178

Problema 6.2

Tasa de llegada	45	45	45
Tasa de servicio	50	50	50
Número de servidores	1	2	3
Factor de utilización	0,90	0,28	0,07
Probabilidad de 0 en el sistema	0,10	0,38	0,40
Número medio en la cola	8,1	0,23	0,03
Tiempo medio en la cola	0,18	0,01	0,01
Número medio en el sistema	9,0	1,13	0,93
Tiempo medio en el sistema	0,2	0,03	0,02
Coste del tiempo de espera	27	0,5	0,28
Coste de los farmacéuticos	10	20	30
Coste total	37	20,5	30,28

Problema 6.3

Aplicando las fórmulas con diversos valores para el número de enfermeras, tendremos que el número óptimo de enfermeras es igual a seis.

Problema 6.4

El número total máximo en el sistema es igual a 4 (tres en cola y uno siendo atendido). Si aplicamos las fórmulas correspondientes (ver sección 6.6), la proporción de personas perdidas porque el sistema está lleno es igual a 18,5 %. Es decir, que en media llegan 8 personas por hora, en cuatro horas llegarán en promedio 32 personas, de las cuales se perderán 9,4 pacientes, con un coste igual a 188 euros. Ahora bien, si aplicamos la política de vales, vendrán en promedio 5 personas por hora, o sea 20 pacientes en cuatro horas y los doce restantes vendrán por la mañana (si consideramos que todos aceptan los vales). Por lo tanto, el coste por vale tiene que ser, como máximo igual a 188/12 = 15,6 euros, ya que si este coste es mayor valdrá más la pena perder pacientes que implantar la política de vales.

8. BIBLIOGRAFÍA

En este apartado damos algunas referencias de libros que pueden ser muy útiles para complementar los temas abordados en este texto. Los asteriscos entre paréntesis indican el grado de dificultad matemática.

Bazaraa, M. (1994): «Programación Lineal y flujo en redes», Limusa (***).

Bronson, R. (1992): «Investigación de Operaciones», McGraw Hill (**).

Daskin, M. (1995): «Network and Discrete Location», Wiley Interscience (****).

Drezner, Z., editor (1995): «Facility Location: A survey of Applications and Methods», Springer (*****).

Eppen, G. y Gould, F. (1984): «Investigación de Operaciones en la Ciencia Administrativa», Prentice Hall (**).

Fernández, R. y Castrodeza, C. (1989): «Programación Lineal», Ariel Economía (**).

Guerras, L. (1989): «Gestión de Empresas y Programación Multicriterio», Esic (**).

Hillier, F. y Lieberman, G. (1997): «Introducción a la Investigación Operativa», McGraw Hill (***).

Puerto, J., editor (1994): «Lecturas en Teoría de Localización», Universidad de Sevilla (*****).

Romero, C. (1993): «Teoría de la Decisión Multicriterio: Conceptos, Técnicas y Aplicaciones», Alianza Universidad (**).

Prawda, J. (1990): «Métodos y Modelos de Investigación de Operaciones», Limusa (***).

Taha, H. (1991): «Investigación de Operaciones», Alfaomega (***).

Villalba, D. y Jerez, M. (1990): «Sistemas de Optimización para la Planificación y Toma de Decisiones», Pirámide (**).

Desde que inició sus actividades, la Fundación BBV ha sido la respuesta institucional del Grupo BBV a la voluntad y al compromiso de complementar una sólida estrategia económica y financiera de su gestión con un firme programa de sensibilidad social y de creación cultural, orientados a la mejora del entorno en el que desarrolla su actividad.

La Fundación BBV trabaja en proyectos de investigación para detectar los problemas que más afectan a la sociedad española, a través de los estudios multidisciplinares, la reflexión y el debate posterior. Pretende, incluso, ir más allá, analizando los problemas desde el contexto europeo y desde la perspectiva internacional.

En sus cerca de diez años de existencia, la Fundación BBV ha desarrollado rigurosas investigaciones sobre cuestiones como la sanidad y la salud, el Estado de Bienestar, el futuro del trabajo y la inactividad laboral, la ética financiera o la movilidad urbana. Es importante destacar otro conjunto de estudios llevados a cabo por la Fundación, tales como la identidad cultural y nacional y el nuevo orden mundial, las alternativas a la sociedad competitiva, y Visiones de Europa.

La Fundación ha hecho también una notable contribución al campo del conocimiento económico y de la realidad social, con investigaciones sobre magnitudes como el *stock* de capital, inversión, renta, producción, etc., de España y sus provincias y comunidades, con datos que cubren ya los últimos cuarenta años de la economía española.

En resumen, hasta el momento la Fundación BBV ha organizado 170 encuentros, realizado cerca de 500 investigaciones, 270 seminarios y más de 450 conferencias. Su Programa Cátedra ha posibilitado la estancia y trabajo en centros españoles de científicos extranjeros de más de 40 universidades, y de científicos españoles en la Universidad de Cambridge; y se sitúa ya en 4.250 su red de colaboradores.

La Fundación BBV mantiene el compromiso de dar a conocer a la sociedad los resultados alcanzados en el marco de sus proyectos y actividades. Documenta, centro editorial de la Fundación, tiene como misión la edición de las publicaciones derivadas de las actuaciones de la Fundación BBV, cuyo catálogo de publicaciones contiene 175 títulos.

FUNDACION BBV

Debido a la creciente complejidad de la gestión y administración de las organizaciones, en las últimas décadas se han desarrollado una serie de métodos cuantitativos destinados a ayudar a los gestores a tomar decisiones. Este libro presenta algunos de los métodos más relevantes utilizados actualmente, como son la programación lineal y la gestión de colas. El nivel de complejidad matemática se mantiene al mínimo nivel posible, y se hace especial énfasis en el planteamiento de modelos y en explicar cómo algunas de las técnicas existentes pueden ayudar a solucionar problemas que aparecen en cualquier organización. Por otro lado, se hace hincapié en la formulación de problemas que aparecen en la gestión de instituciones sanitarias.

La Fundación BBV pretende que este libro sirva básicamente para dos tipos de usuarios: el administrador general o administrador en potencia, que pueden sacar provecho de su uso y de la comprensión de los modelos cuantitativos; y para estudiantes de cursos de posgrado o masters en gestión y administración.

ISBN 84-95163-25-X
9 788495 163257

MÉTODOS CUANTITATIVOS PARA LA TOMA DE DECISIONES

FUNDACION BBV