GS.TSKH. Phan Quốc Khánh – TS. Trần Huệ Nương

Quy hoạch tuyến tính

GIÁO TRÌNH HOÀN CHỈNH:

Lí thuyết cơ bản, Phương pháp đơn hình Bài toán mạng, Thuật toán điểm trong

(Tái bản lần thứ hai)

Nhà xuất bản Giáo dục

Quyển sách này là kết quả kinh nghiệm nhiều năm giảng dạy QUY HOẠCH TUYẾN TÍNH, được biên soạn trong khuôn khổ của chương trình *"Tối ưu hóa - Toán ứng dụng"* theo hợp tác khoa học giữa *Hội đồng Liên trường Đại học nói tiếng Pháp của Vương quốc Bỉ* và *Đại học Quốc gia Việt Nam - Thành phố Hồ Chí Minh*, do *Giáo sư Nguyễn Văn Hiền*, đại học Namur, chủ nhiệm chương trình, gợi ý và tạo điều kiện thực hiện. Các tác giả rất cảm ơn C.I.U.F. - C.U.D. / C.U.I. đã cho phép viết sách này trong chương trình "Tối ưu hóa - Toán ứng dụng".

Trong quá trình biên soạn, các tác giả đã tham khảo nhiều tài liệu mới nhất, trong đó có hầu hết các sách về quy hoạch tuyến tính do các nhà xuất bản nổi tiếng thế giới phát hành từ 1995 đến 1998, và đã tranh thủ được kinh nghiệm của nhiều giáo sư - đồng nghiệp Bỉ.

CÁC TÁC GIẢ

Cet ouvrage est d'abord le fruit d'une longue expérience d'enseignement de la PROGRAMATION LINÉAIRE. Il a été préparé dans le cadre du Programme d'Activités "*Optimisation - Mathématiques Appliquées*" du Projet d'Asie du Sud - Est/Vietnam de la Coopération Universitaire au Développement du *Conseil Interuniversitaire de la Communauté Française de Belgique* (C.I.U.F. - C.U.D. / C.U.I.) et de *l'Université Nationale du Vietnam à Hochiminh - Ville*. La rédaction de ce livre a bénéficié de la stimulation intellectuelle et de l'encouragement moral du *Professeur Nguyen Van Hien* des Facultés Universitaire Notre - Dame de la Paix de Namur, Responsable du Programme d'Activités. Les auteurs tiennent à lui témoigner leur profonde gratitude aussi qu'envers le C.I.U.F. - C.U.D. / C.U.I. qui leur a permis de mener à bien cette tâche.

Lors de la rédaction de ce manuel, les auteurs ont pu consulter la littérature récente sur le sujet et, en particulier, les livres publiés entre 1995 - 1998 par les maisons d'édition de renommée internationalle. Ils ont pu également bénéficier des expériences des collègues belges.

LỜI NÓI ĐẦU

Cuối tháng 8 - 1997 hơn hai ngàn nhà khoa học từ khắp các nước, tham dự Hội nghị Quốc tế **"Quy hoạch Toán học"** tại Lausanne, Thụy Sỹ, đã làm lễ kỉ niệm 50 năm ngày phương pháp đơn hình của Dansic (George Bernard Dantzig, 1914-) được công bố, cũng được coi là ngày QUY HOẠCH TUYẾN TÍNH chính thức chào đời. Tham gia hội nghị và nhất là được nghe các bài nói của giáo sư Dantzig tại đây, chúng tôi mới thấy thật rõ vai trò lớn lao của phương pháp đơn hình, sự phát triển như vũ bão, nhất là trong thập niên vừa qua, của quy hoạch tuyến tính và mối quan hệ của phương pháp đơn hình với phương pháp ellipsoid của Khachian 1979 và các phương pháp điểm trong mở đầu bằng phương pháp biến đổi chiếu nổi tiếng của Karmarkar 1984. (Mặc dù trước đây chúng tôi đã gặp giáo sư Dantzig vài lần.)

Nhìn lại trong nước, chúng tôi thấy mừng là Quy hoạch tuyến tính được giảng dạy ở khắp các trường đại học và gần như cho tất cả các ngành. Nhưng bên cạnh điểm chung này với các nước khác, đáng tiếc là giáo trình và tài liệu của chúng ta khá lạc hậu. Nhiều nơi vẫn dạy theo giáo trình từ các tài liệu xuất bản trong thập niên sáu bảy mươi. Các phương pháp nói trên không được đề cập đến, trừ phương pháp đơn hình. Một sự phù hợp may mắn là đầu năm 1997, Khoa Toán - Tin học thuộc Trường Đại học Khoa học Tự nhiên TP Hồ Chí Minh (Đại học Tổng hợp trước đây) đã bắt đầu thảo luận và chuẩn bị cải tiến chương trình giảng dạy tất cả các chuyên ngành. Chúng tôi đang biên soạn lại giáo trình quy hoạch tuyến tính, thì được Giáo sư Nguyễn Văn Hiền, thuộc Đại học Namur, Vương quốc Bỉ gợi ý viết sách này trong chương trình **"Tối ưu hóa - Toán ứng dụng"** do Giáo sư là chủ nhiệm phía Bỉ và tôi là chủ nhiệm phía Việt Nam.

Giáo trình **QUY HOẠCH TUYẾN** *TÍNH* này trình bày cơ sở lí thuyết, các phương pháp chính và các bài toán cơ bản của quy hoạch tuyến tính, theo dạng mới nhất hay được vận dụng ở các sách nước ngoài xuất bản trong thời gian 1995 - 98. Nội dung sách là sự mở rộng của một giáo trình cơ sở 45 tiết và một chuyên đề nâng cao 60 tiết, mà chúng tôi đã giảng dạy nhiều năm ở một số trường đại học.

Sách được viết ở dạng thuận lợi cho độc giả rộng rãi, kể cả sinh viên tự học. Để hiểu, độc giả chỉ cần có kiến thức đại số tuyến tính và phép tính vi phân ở chương trình năm thứ nhất của bất kì đại học nào.

Những nội dung cần thiết, rộng hơn và sâu hơn, về quy hoạch tuyến tính cũng như hướng dẫn bài tập, một phần không thể tách rời của chính ngay hai giáo trình cho sinh viên nói trên, chúng tôi sẽ trình bày trong quyển sách tiếp theo **"QUY HOẠCH TUYẾN TÍNH. KIẾN THỨC BỔ SUNG VÀ BÀI TẬP"**.

Phân công biên soạn trong giáo trình này như sau:

- Trần Huệ Nương *viết các Chương 1, 2, 3, 6, 7, 8, 9;*

- Phan Quốc Khánh *(chủ biên) viết các Chương 4, 5, 10, 11, 12, 13, 14, 15.*

Gần toàn bộ việc sưu tầm tài liệu, xây dựng đề cương và việc biên soạn lần đầu của một phần quyển **QUY HOẠCH TUYẾN TÍNH** *này được hoàn thành trong thời gian chúng tôi làm việc ở Đại học Namur theo chương trình* "Tối ưu hóa - Toán ứng dụng" *nói trên. Chúng tôi rất biết ơn Khoa Toán, đại học Namur, đã giúp chúng tôi mọi điều kiện thuận lợi. Sự tri ân đặc biệt được dành cho Giáo sư Nguyễn Văn Hiền, với vai trò chủ nhiệm chương trình và sự giúp đỡ chuyên môn của Giáo sư, cùng các giáo sư Jean - Jacques Strodiot (Đại học Namur) và Etienne Loute (Đại học Saint Louis). Các tác giả cũng chân thành cảm ơn Hội đồng liên trường đại học nói tiếng Pháp của Vương quốc Bỉ, Đại học quốc gia TP Hồ chí Minh và trường Đại học Khoa học tự nhiên đã cho phép và tạo điều kiện thuận lợi để việc biên soạn này thuộc chương trình hợp tác Bỉ - Việt Nam. Lòng biết ơn thường xuyên nhất là đối với các đồng nghiệp thuộc Khoa Toán - Tin học, đã cộng tác trong chương trình giảng dạy của chúng tôi nhiều năm. Chúng tôi cũng rất cảm ơn em Phan Khánh Linh đã cung cấp tài liệu về quy hoạch tuyến tính ở các trường đại học Australia.*

Sự biết ơn cũng dành cho TS. Nguyễn Cao Thắng - Biên tập viên và Nhà Xuất bản Giáo dục đã kịp thời phát hành sách này ngay đầu năm 2000. Đặc biệt, năm 2000 là năm đầu tiên UNESCO chọn làm năm cho một ngành khoa học. Là một ngành khoa học có vai trò cực kì quan trọng xuyên suốt lịch sử văn minh của loài người, với vị trí của mình trong giáo dục, với tác động quyết định của mình với nhiều ngành khoa học khác (Cơ học, Vật lí, Công nghệ thông tin, ...) và với ngôn ngữ phổ dụng của mình, TOÁN HỌC đã được UNESCO chọn làm **Năm Toán học Thế giới 2000** *(Mathematical World Year 2000), như các bạn thấy logo ở bìa 4 của cuốn sách này. Cuối cùng, sự tri ân của tác giả luôn dành cho những ý kiến phê bình, sửa đổi, cải tiến sách, luôn được mong đợi từ độc giả.*

TP Hồ Chí Minh, ngày 1-12-1999

PHAN QUỐC KHÁNH

MỞ ĐẦU

Loài người đã biết chọn "phương án" hành động trong các công việc của mình từ thời cổ đại. Đầu tiên người ta biết chọn phương án "chấp nhận được", theo các tiêu chuẩn từ mức độ cảm tính đến có cơ sở khoa học và định lượng. Khi có nhiều phương án chấp nhận được, điều mong muốn tự nhiên là chọn cái tốt nhất (tức là "tối ưu"), cũng lại theo một hoặc một số tiêu chuẩn nào đó. Dần dần người ta biết "mô hình hóa toán học" công việc của mình, tức là diễn đạt công việc đó ở dạng phương trình toán học, diễn giải thế nào là tiêu chuẩn chấp nhận được, tiêu chuẩn tối ưu...

Cùng với sự hình thành khá hoàn chỉnh của phép tính vi tích phân vào thế kỉ 17, "bài toán cực trị" trong lí thuyết này là một mô tả toán học chính xác của bài toán tối ưu.

Đồng thời, trong vài thế kỉ vừa qua, các bài toán về kinh tế và quản lí cũng dẫn người ta đến lí thuyết tối ưu từ một góc độ khác.

Nhưng lí thuyết toán học về tối ưu chỉ được hình thành và phát triển mạnh như một lĩnh vực khoa học quan trọng từ khoảng giữa thế kỉ này. Tùy theo dạng các bài toán được nghiên cứu, đặc điểm của mô hình toán học và công cụ xét chúng hoặc phạm vi áp dụng..., nhiều lĩnh vực khá gần nhau và đan xen với nhau của lí thuyết được hình thành với các tên gọi khác nhau: *tối ưu hóa* (optimization), *quy hoạch toán học* (mathematical programming), *vận trù học* (operation research), *điều khiển tối ưu* (optimal control), *lí thuyết các bài toán cực trị* (theory of extremal problems) *phép tính biến phân* (variational calculus)...

Trong quy hoạch toán học lại có quy hoạch tuyến tính, quy hoạch phi tuyến và nhiều lĩnh vực đặc thù khác như quy hoạch nguyên, quy hoạch phân thức, quy hoạch động, quy hoạch ngẫu nhiên...

Vậy *quy hoạch tuyến tính* (linear programming) là gì? Có thể tạm định nghĩa quy hoạch tuyến tính là lĩnh vực toán học nghiên cứu các bài toán tối ưu trên hữu hạn biến mà hàm mục tiêu và các ràng buộc đều là hàm và các phương trình hoặc bất phương trình tuyến tính. Đây chỉ là một định nghĩa mơ hồ. Chúng ta sẽ xác định cụ thể bài toán quy hoạch tuyến tính ngay trong chương này.

Quy hoạch tuyến tính được coi là ra đời vào năm 1947, khi Dantzig công bố phương pháp đơn hình để giải các bài toán xuất phát từ việc lập kế hoạch cho không quân Mỹ. Vậy có thể nói là, cũng như phép tính vi tích phân hình thành vào thế kỉ 17 từ việc giải các bài toán cơ học, quy hoạch tuyến tính hình thành vào giữa thế kỉ 20 do nhu cầu của các bài toán quản lí. Nhưng ngay sau đó nó đã được phát triển rất nhanh nhờ sự phát hiện ra rằng nhiều bài toán, có thể khác xa nhau về hình thức, được đưa về quy hoạch tuyến tính và giải bằng phương pháp đơn hình. Cũng ngay trong năm 1947, T.C. Koopmans đã chỉ ra rằng quy hoạch tuyến tính là công cụ tuyệt vời để phân tích lí thuyết kinh tế cổ điển, chẳng hạn mô hình của L. Walras đề xuất từ 1874.

Cũng như mọi lĩnh vực khoa học khác, QUY HOẠCH TUYẾN TÍNH được sinh ra không phải bất ngờ sau vài ngày, hoặc vài tháng. Lí thuyết toán học chính và là chỗ dựa căn bản để Dantzig phát minh ra phương pháp đơn hình là lí thuyết về các bất đẳng thức tuyến tính, đã được nghiên cứu kĩ bởi Fourier từ 1826.

Dantzig được coi là cha đẻ của QUY HOẠCH TUYẾN TÍNH có thể cũng chỉ vì quyển sách sớm hơn nhiều "Các phương pháp toán học trong tổ chức và kế hoạch hóa sản xuất" của L.V. Kantorovich do nhà xuất bản Đại học quốc gia Leningrad in năm 1939, ít được biết đến thậm chí ở Nga và chỉ được dịch sang tiếng Anh ở Mỹ năm 1960. (Có thể 1939-47 chính là những năm đại chiến thế giới thứ hai và ảnh hưởng của nó!?) Trong quyển sách rất giá trị này Kantorovich đã nêu bật vai trò của một lớp bài toán quy hoạch tuyến tính và đề xuất thuật toán sơ bộ để giải chúng. Tuy bị biết đến muộn màng, ông được coi là một nhà toán học hàng đầu trong lịch sử thế giới về toán kinh tế. Năm 1975 ông cùng Koopmans được Viện Hàn lâm Hoàng gia Thụy Điển trao giải thưởng Nobel về khoa học kinh tế.

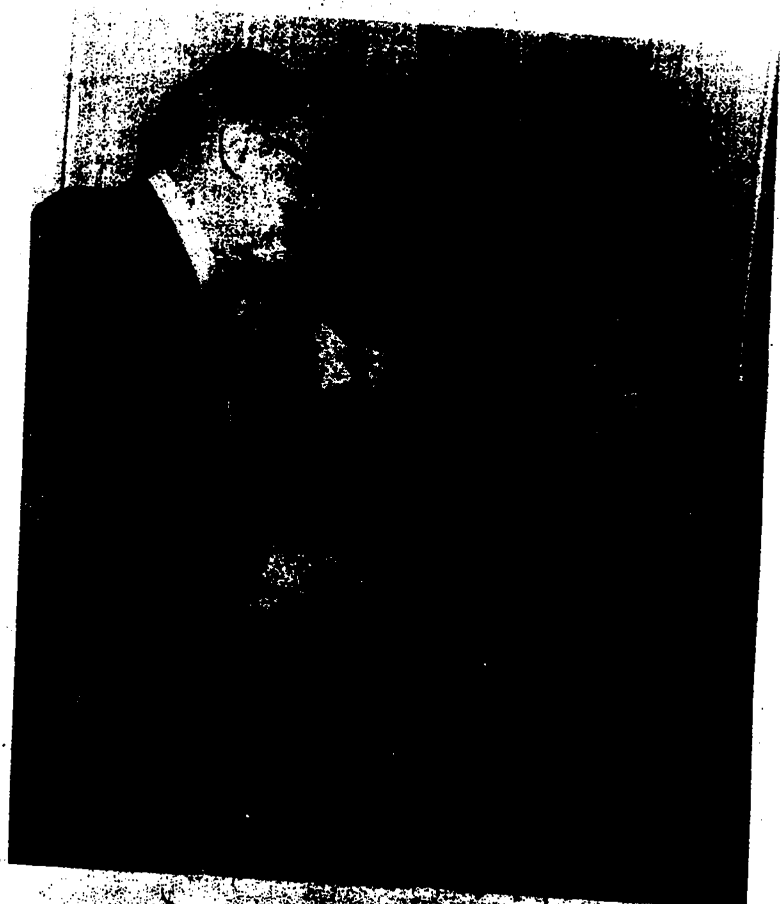
Từ khi ra đời, phương pháp đơn hình gần như thống trị trong QUY HOẠCH TUYẾN TÍNH. Dựa trên nhận xét rằng miền chấp nhận được của QUY HOẠCH TUYẾN TÍNH là tập lồi đa diện và nếu bài toán đã có nghiệm tối ưu thì nó luôn có nghiệm tối ưu là đỉnh của tập lồi đa diện này, nội dung của phương pháp đơn hình là cố gắng tìm một đỉnh chấp nhận được rồi mỗi bước là chuyển từ một đỉnh chấp nhận được sang đỉnh kề (tức là có một cạnh nối nhau) tốt hơn (tức là có giá trị hàm mục tiêu tương ứng tốt hơn, chẳng hạn trong bài toán tìm maximum, là lớn hơn). Vì số đỉnh là hữu hạn, thuật toán sẽ kết thúc ở đỉnh tối ưu sau hữu hạn bước. Phương pháp này tận dụng triệt để cấu trúc tuyến tính, khác hẳn các phương pháp tối ưu phi tuyến.

Đối với các bài toán cỡ lớn (có thể đến chục nghìn biến và mấy trăm ràng buộc) phải dùng đến máy tính, phương pháp đơn hình cũng được kiểm nghiệm qua mấy chục năm áp dụng là rất hiệu quả, với thời gian tính toán khá ngắn. Có một lĩnh vực rất quan trọng nghiên cứu về độ phức tạp tính toán. Ở đây hiệu quả của một phương pháp được đo bằng thời gian tính toán phụ thuộc theo cỡ của bài toán. Cỡ của bài toán là thời gian cần thiết để máy tính ghi các dữ liệu của bài toán (do đó có thể nói nôm na là cỡ của bài toán đặc trưng bởi số biến và số ràng buộc). Một thuật toán được gọi là có độ phức tạp đa thức, hoặc gọi tắt là *thuật toán đa thức* (polynomial algorithm) nếu thời gian tính toán phụ thuộc theo quy luật đa thức vào cỡ của bài toán, kể cả trong trường hợp xấu nhất.

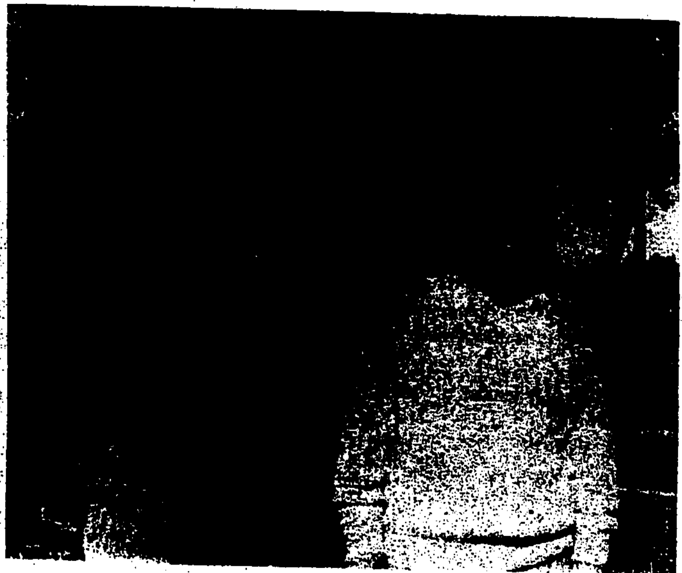
Năm 1972, V.L.Klee và G.J.Minty đã đưa ra thí dụ một QUY HOẠCH TUYẾN TÍNH mà thời gian tính toán theo thuật toán đơn hình là hàm mũ của cỡ bài toán. Vậy thật tiếc là phương pháp đơn hình không phải là một thuật toán đa thức. Những cố gắng tìm thuật toán đa thức cho QUY HOẠCH TUYẾN TÍNH đã dẫn đến thời điểm kịch tính là năm 1979 khi L.G.Khachian đưa ra thuật toán ellipsoid có độ phức tạp đa thức làm chấn động dư luận toán học thế giới trong vài năm. Nhưng rồi người ta thấy rằng phương pháp này chỉ có ý nghĩa lí thuyết vì qua nhiều áp dụng thực tế nó đều kém phương pháp đơn hình về thời gian tính toán. Vì theo quan điểm áp dụng thì "thời gian trung bình" cho đa số các bài toán hay gặp quan trọng hơn thời gian ứng với trường hợp xấu nhất về lí thuyết gần như không gặp trong thực tế.

Bước ngoặt tiếp theo trong lịch sử phát triển của QUY HOẠCH TUYẾN TÍNH còn lớn hơn, đó là năm 1984 khi nhà toán học Mỹ N.K.Karmarkar công bố một phương pháp điểm trong có độ phức tạp đa thức. Khác hẳn phương pháp đơn hình, xây dựng dãy các điểm biên (là đỉnh) của miền chấp nhận được, phương pháp Karmarkar xây dựng dãy các điểm trong sao cho hội tụ về điểm biên là nghiệm tối ưu. Tư tưởng này đã kích thích các nhà toán học đề xuất nhiều phương pháp điểm trong khác sau đó. Các phương pháp này được chứng tỏ qua một số bài toán cỡ lớn có tính "suy biến" là nhanh hơn phương pháp đơn hình (nhưng áp dụng phương pháp đơn hình thường vẫn tiện lợi hơn). Vì vậy Karmarkar là một trong các tên tuổi toán học được nhắc đến nhiều nhất trong thập niên vừa qua. Nhưng lại một lần nữa, tác giả thực sự của bước ngoặt này, mà sớm hơn nhiều, là I.I.Dikin, vì ông đã công bố phương pháp thực chất là trùng với phương pháp Karmarkar ở tạp chí "Báo cáo của Viện Hàn lâm khoa học Liên Xô", từ năm 1967, nhưng chưa chú ý đến độ phức tạp tính toán và bài báo đã không được để ý đến.

Khác với các sách QUY HOẠCH TUYẾN TÍNH đã xuất bản trong nước, ở đây chúng tôi trình bày cả phương pháp đơn hình và các dạng cải tiến, mở rộng của nó (thường gọi chung là các phương pháp biên), các phương pháp điểm trong và phương pháp ellipsoid. Các bài toán về mạng và dòng cũng được giới thiệu. Như vậy, tất cả các nội dung cơ bản nhất của QUY HOẠCH TUYẾN TÍNH đều được bao hàm trong sách này.



Koopmans, Dantzig, Kantorovich (Laxenburg, 1976)



Dantzig, Khachian (Asilomar, 1990)

Karmarkar, Grötschel, Iri (Tokyo, 1988)

PHẦN I

LÍ THUYẾT CƠ BẢN VÀ PHƯƠNG PHÁP ĐƠN HÌNH

CHƯƠNG 1

QUY HOẠCH TUYẾN TÍNH VÀ CÁCH GIẢI ĐƠN GIẢN

1.1 BÀI TOÁN QUY HOẠCH TUYẾN TÍNH TRONG THỰC TẾ

Trước khi định nghĩa quy hoạch tuyến tính và nghiên cứu nó, ta xét một vài bài toán thực tế điển hình (và đơn giản) có thể phát biểu toán học thành quy hoạch tuyến tính. Chúng ta sẽ thấy mô hình toán học thật đẹp và tự nhiên, đơn giản làm sao. Thế nhưng mãi đến năm 1947, G. B. Dantzig mới đưa ra được mô hình toán học này khi nghiên cứu các bài toán lập kế hoạch cho không quân Mỹ. Lúc đầu ông gọi là "Quy hoạch trong cấu trúc tuyến tính" (Programming in a linear structure). Hè năm 1948, khi Tjalling Koopmans cùng Dantzig đi dạo trên bãi biển Santa Monica ở Los Angeles, Koopmans nói "Sao không gọi '*quy hoạch trong cấu trúc tuyến tính*' ngắn gọn là '*quy hoạch tuyến tính*'". Dantzig đồng ý ngay "Đúng vậy! từ nay sẽ là '*quy hoạch tuyến tính*'" ('linear programming'). Nhiều năm sau, Albert Tucker dùng tên ngắn hơn là linear program. Ngay sau khi Dantzig đưa ra quy hoạch tuyến tính, người ta thấy rất nhiều bài toán thực tế thuộc các lĩnh vực khác nhau có thể mô tả toán học là quy hoạch tuyến tính. Mô hình này vẫn thống trị cho đến nay. Để thấy rõ mô hình toán học là quan trọng nhường nào xin nói thêm rằng hiện đang có một bài toán mở chưa ai trả lời được là: "Liệu có một lớp quy hoạch tuyến tính hẹp hơn mà vẫn mô tả được (gần như) tất cả các bài toán thực tế?" Đây không phải là câu hỏi đơn giản được đặt ra theo logic hình thức của vấn đề. Cơ sở sâu xa của nó là ở sự so sánh bất phân thắng bại giữa phương pháp đơn hình và phương pháp điểm trong trong thực tế sử dụng, mặc dù người ta đã chứng minh được rằng trong trường hợp xấu nhất các phương pháp điểm trong vẫn hội tụ với thời gian là đa thức theo "số đo" dữ liệu của bài toán (xem Chương 13 - 15). Còn với phương pháp đơn hình thì thời gian này là hàm mũ theo số đo dữ liệu, theo thí dụ của Victor L. Klee và George J. Minty (xem Mục 5.2).

Vậy phải chăng thực tế chỉ cần một lớp quy hoạch tuyến tính hẹp hơn (cũng tức là một mô hình toán học hơi khác), mà với nó thì phương pháp đơn hình có thời gian đa thức, tức là thời gian tính toán cho trường hợp xấu nhất vẫn là đa thức của số đo dữ liệu.

THÍ DỤ 1.1.1 *Bài toán kế hoạch sản xuất* - (Production planning problem). Công ty Reddy Mikks sản xuất sơn nội thất và sơn ngoài trời. Nguyên liệu gồm 2 loại A và B với trữ lượng là 6 tấn và 8 tấn tương ứng. Để sản xuất một tấn sơn nội thất cần 2 tấn nguyên liệu A và 1 tấn nguyên liệu B. Hai số tương ứng của sơn ngoài trời là 1 tấn và 2 tấn. Qua tiếp thị được biết nhu cầu thị trường là như sau (cho một ngày):

- Nhu cầu sơn nội thất không hơn nhu cầu sơn ngoài trời quá 1 tấn;

- Nhu cầu cực đại của sơn nội thất là 2 tấn.

Giá bán sỉ là 2000 USD một tấn sơn nội thất và 3000 USD một tấn sơn ngoài trời.

Vấn đề là cần sản xuất mỗi ngày như thế nào để doanh thu là lớn nhất.

Gọi x_1 và x_2 là số lượng (tính theo tấn) sơn ngoài trời và sơn nội thất tương ứng cần sản xuất trong ngày. Đây sẽ là các *biến* (variable) hoặc *phương án* (alternative) của bài toán. Khi đó doanh thu trong ngày sẽ là

$$z = 3x_1 + 2x_2$$

và được gọi là *hàm mục tiêu* (objective function).

Các ràng buộc trên biến x_1, x_2, sẽ gọi là các *ràng buộc* (constraint) của bài toán, là như sau. Nguyên liệu sử dụng không được quá trữ lượng:

$$x_1 + 2x_2 \le 6 \quad \text{(nguyên liệu A)},$$
$$2x_1 + x_2 \le 8 \quad \text{(nguyên liệu B)}.$$

Sản xuất không nhiều hơn nhu cầu thị trường:

$$x_2 - x_1 \le 1,$$
$$x_2 \le 2.$$

Sản lượng phải là số thực không âm:

$$x_1 \ge 0,\ x_2 \ge 0.$$

Ta gọi phương án (x_1, x_2) là *chấp nhận được* (feasible) nếu nó thỏa mọi ràng buộc. Khi đó (x_1, x_2) cũng gọi là *nghiệm chấp nhận được* (feasible solution).

Vậy bài toán trở thành:

Tìm phương án chấp nhận được làm cực đại hàm mục tiêu z và được viết ở dạng toán học như sau

$$\begin{aligned}
\max z = 3x_1 + 2x_2&,\\
x_1 + 2x_2 &\le 6,\\
2x_1 + x_2 &\le 8,\\
-x_1 + x_2 &\le 1,\\
x_2 &\le 2,\\
x_1 \ge 0,\quad x_2 &\ge 0.
\end{aligned}$$

THÍ DỤ 1.1.2 *Bài toán khẩu phần ăn* - (Diet problem). Giả sử người ta muốn chế biến món ăn từ nhiều thành phần (thực phẩm) sao cho đủ các chất bổ (như chất đạm, chất béo, chất đường...) mà giá thành lại rẻ nhất.

Giả sử có n thành phần, với giá một đơn vị (khối lượng) thành phần j là c_j, j = 1, ..., n. Đồng thời có m chất. Biết rằng một đơn vị thành phần j chứa a_{ij} đơn vị chất i, i = 1, ..., m, và mức chấp nhận được số đơn vị chất i trong hỗn hợp là nằm giữa $l_i \geq 0$ và $u_i \geq 0$, i = 1, ..., m.

Gọi x_j là số đơn vị khối lượng của thành phần j trong một đơn vị khối lượng của món ăn. Khi đó, tương tự như thí dụ trên, dễ thấy bài toán trở thành

$$\min z = \sum_{j=1}^{n} c_j x_j ,$$

$$l_i \leq \sum_{j=1}^{n} a_{ij} x_j \leq u_i,\ i = 1, ..., m, \tag{1.1}$$

$$\sum_{j=1}^{n} x_j = 1,$$

$$x_j \geq 0,\ j = 1, ..., n.$$

THÍ DỤ 1.1.3. *Bài toán vận tải* - (Transportation problem). Hàng hóa được vận chuyển từ m kho đến n cửa hiệu bán lẻ. Lượng hàng ở kho i là $s_i \geq 0$ (tấn), i = 1, ..., m và cửa hiệu j có nhu cầu $d_j \geq 0$ (tấn), j = 1, ..., n. Cước vận chuyển một tấn hàng từ kho i đến cửa hiệu j là c_{ij} đồng. Giả sử tổng hàng ở các kho và tổng nhu cầu bằng nhau:

$$\sum_{i=1}^{m} s_i = \sum_{j=1}^{n} d_j . \tag{1.2}$$

Bài toán đặt ra là lập kế hoạch vận chuyển để tiền cước là nhỏ nhất, với điều kiện là mỗi cửa hàng đều nhận đủ và mỗi kho đều trao hết hàng.

Gọi lượng hàng vận chuyển từ kho i đến cửa hàng j là x_{ij}, thì kế hoạch vận chuyển, tức là phương án theo định nghĩa chung, là ma trận (x_{ij}) cấp m x n. Dạng toán học của bài toán là

$$\min z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij},$$

$$\sum_{j=1}^{n} x_{ij} = s_i, \ i = 1, ..., m,$$

$$\sum_{i=1}^{m} x_{ij} = d_j, \ j = 1, ..., n, \qquad (1.3)$$

$$x_{ij} \geq 0, \ i = 1, ..., m, \ j = 1, ..., n.$$

Mô hình này gọi là *mô hình vận tải đóng* (closed transportation model). Nếu không có giả thiết (1.2) và ràng buộc (1.3) đổi lại là $\sum_{j=1}^{n} x_{ij} \leq s_i$, tức là các kho có thể không trao hết, thì mô hình được gọi là *mô hình vận tải mở* (open transportation model).

1.2 ĐỊNH NGHĨA QUY HOẠCH TUYẾN TÍNH

Các thí dụ trên thuộc các lĩnh vực khác nhau nhưng được đưa về dạng toán học có nét chung là:

Phải xác định các *biến quyết định* (decision variable), gọi tắt là *biến*, hoặc *phương án* (alternative), thỏa mãn các ràng buộc sao cho làm cực đại hoặc cực tiểu hàm mục tiêu. Hơn nữa cả hàm mục tiêu và các ràng buộc đều tuyến tính (bậc nhất) theo biến quyết định.

Nhận xét thêm là việc tìm cực đại có thể dễ dàng chuyển thành tìm cực tiểu và ngược lại, vì max z = – min (–z). Do đó các thí dụ trên đều là trường hợp riêng của bài toán

$$\begin{aligned}
&\min z = c^T x, && \\
&a_i^T x \geq b_i, && i \in M_1, \\
&a_i^T x \leq b_i, && i \in M_2, \\
&a_i^T x = b_i, && i \in M_3, \\
&x_j \geq 0, && j \in N_1, \\
&x_j \leq 0, && j \in N_2,
\end{aligned} \tag{1.4}$$

ở đây M_1, M_2, M_3, N_1 và N_2 là tập hợp chỉ số nào đó, c^T là chuyển vị (transpose) của vectơ c. x, a_i và c là các vectơ n thành phần. b_i là các số thực. Ta luôn quy ước vectơ là vectơ cột, vậy c^T là vectơ hàng. Chú ý là ràng buộc kiểu (1.1) đã được tách thành hai ràng buộc để đưa về dạng trên đây. $\min z = c^T x$ cũng thường viết gọn là $\min c^T x$.

Ràng buộc đẳng thức có thể viết tương đương ở dạng bất đẳng thức như sau:

$$a_i^T x = b_i \Leftrightarrow a_i^T x \leq b_i,\ a_i^T x \geq b_i.$$

Ràng buộc dấu $x_j \geq 0$ là trường hợp riêng của ràng buộc bất đẳng thức dạng $a^T x \geq b$ với $a^T = (0, ..., 0, \underset{j}{\underline{1}}, 0, ..., 0)$ và $b = 0$. Ngoài ra, rõ ràng $a_i^T x \leq b_i$ tương đương với $-a_i^T x \geq -b_i$.

Giả sử sau khi thực hiện các biến đổi trên (1.4) có tất cả m ràng buộc dạng $a_i^T x \leq b_i$. Ký hiệu A là ma trận cấp m x n có các hàng là a_i^T và $b = (b_1, ..., b_m)^T$, thì bài toán (1.4) có dạng ma trận là

$$\begin{aligned}
&\min c^T x, \\
&Ax \leq b.
\end{aligned} \tag{1.5}$$

Ở đây bất đẳng thức vectơ $Ax \leq b$ được hiểu là m bất đẳng thức theo thành phần, tức là $a_i^T x \leq b_i$, $i = 1, ..., m$. Vậy quy hoạch tuyến tính là bài toán có dạng tổng quát là (1.4) hoặc viết ở dạng ma trận

(1.5). Dùng lời diễn đạt thì quy hoạch tuyến tính là bài toán tìm cực tiểu (hoặc cực đại) của hàm mục tiêu tuyến tính với các ràng buộc bất đẳng thức và đẳng thức tuyến tính. Tuy vậy, trong nghiên cứu quy hoạch tuyến tính cũng như khi áp dụng nó, người ta thường dùng hai dạng đặc thù

Dạng chính tắc: (canonical form)	$\min c^T x,$ $Ax = b,$ $x \geq 0;$
Dạng chuẩn: (normal form hoặc standard form)	$\min c^T x,$ $Ax \leq b,$ $x \geq 0.$

Ở đây A là ma trận cấp m x n với các hàng là a_i, $i = 1, ..., m$ và các cột là A_j, $j = 1, ..., n$. c và x là vectơ n chiều và b là vectơ m chiều (Ở đây cũng có sự khác nhau về thuật ngữ. Dantzig định nghĩa dạng chính tắc trên đây là standard form và gọi canonical form là trường hợp đặc biệt của nó, mà ta sẽ gặp sau.) Việc viết riêng các ràng buộc dấu $x \geq 0$ là có nhiều thuận lợi, nhất là khi xét về hình học. Nếu ở bài toán có các biến tự do (về dấu không có ràng buộc), giả sử là x_j, thì ta thay x_j bằng hai biến mới x_j^+ và x_j^- bởi liên hệ $x_j = x_j^+ - x_j^-$, $x_j^+ \geq 0$, $x_j^- \geq 0$. Ta đã biết ràng buộc đẳng thức có thể viết tương đương thành ràng buộc bất đẳng thức. Ngược lại, (để đưa về dạng chính tắc) ràng buộc bất đẳng thức có thể chuyển thành đẳng thức bằng cách đưa thêm vào biến bù như sau. Giả sử ta có ràng buộc bất đẳng thức $a_i^T x \leq b_i$. Ta đưa thêm vào biến mới s_i, gọi là *biến bù* (slack variable) như sau:

$$a_i^T x + s_i = b_i,$$

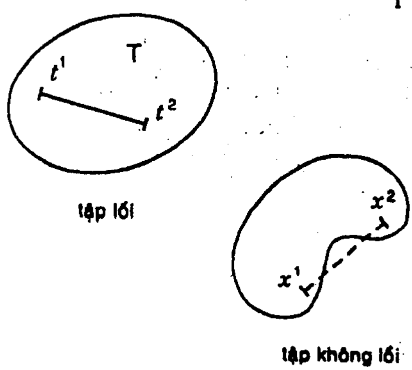
$$s_i \geq 0.$$

(Tương tự, $a_i^T x \geq b_i \Leftrightarrow a_i^T x - s_i = b_i$, $s_i \geq 0$.)

1.3 GIẢI QUY HOẠCH TUYẾN TÍNH ĐƠN GIẢN

1.3.1 Quy hoạch tuyến tính hai biến

Khi bài toán chỉ có hai biến, ta có thể giải quy hoạch tuyến tính bằng hình học dễ dàng. Chú ý rằng trường hợp riêng này cũng cho phép ta tưởng tượng hình học về bài toán tổng quát. Hãy xét thí dụ 1.1 về kế hoạch sản xuất của công ty Reddy Mikks. Nhận xét rằng mỗi ràng buộc bất đẳng thức $a_i^T x \le b_i$ cho ta một *nửa không gian đóng* (closed half space) $\{x \in \mathbf{R}^n : a_i^T x \le b_i\}$ giới hạn bởi *siêu phẳng* (hyperplane) $\{x \in \mathbf{R}^n: a_i^T x = b_i\}$. Trường hợp hai biến, đây là nửa mặt phẳng (đóng) giới hạn bởi đường thẳng. Tập tất cả các nghiệm chấp nhận được gọi là *miền chấp nhận được* (feasible region), hoặc *miền ràng buộc*, là một *tập lồi đa diện* (polyhedron). Tập lồi đa diện là giao của một số hữu hạn các nửa không gian đóng, tức là tập có dạng $\{x \in \mathbf{R}^n: a_i^T x \le b_i,\ i = 1, ..., m\}$; hoặc ở dạng vectơ $\{x \in \mathbf{R}^n: Ax \le b\}$. Chú ý rằng chiều của bất đẳng thức ở đây có thể ở dạng khác. Các tập $\{x \in \mathbf{R}^n: Ax \ge b\}$, $\{x \in \mathbf{R}^n: a_i^T x \le b_i,\ i = 1,...,m;\ a_j^T x \ge b_j,\ j = 1,...,k;\ a_p^T x = b_p,\ p = 1,...,s\}$ cũng là tập lồi đa diện. Nếu tập lồi đa diện giới nội ta gọi nó là *đa diện lồi* (polytope). Trong trường hợp hai biến, đây là đa giác lồi. *Tính lồi* (convexity) là một khái niệm quan trọng. Tập $T \subset \mathbf{R}^n$ gọi là *lồi* (convex) nếu bất kì hai điểm $t^1 \in T$ và $t^2 \in T$ đều có đoạn thẳng nối chúng nằm



H. 1.1

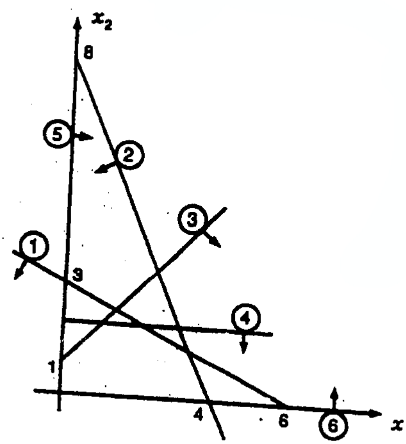
hoàn toàn trong T (xem H.1.1). Chú ý rằng đoạn nối t^1 và t^2 là tập $\{t \in \mathbf{R}^n: t = (1 - \gamma)t^1 + \gamma t^2, 0 \le \gamma \le 1\}$, tức là tập tất cả các *tổ hợp lồi* (convex combination) $(1 - \gamma)t^1 + \gamma t^2$ của t^1 và t^2. Một cách tự nhiên, ta có định nghĩa tổ hợp lồi tổng quát của nhiều điểm. Điểm trong $\mathbf{R}^n$ có dạng $\sum_{i=1}^{m} \gamma_i t^i$, ở đây $\gamma_i \ge 0$ và $\sum_{i=1}^{m} \gamma_i = 1$, gọi là một *tổ hợp lồi* của các điểm t^i, i = 1, ..., m.

Nhắc lại là bài toán công ty Reddy Mikks là quy hoạch tuyến tính

$$\max 3x_1 + 2x_2,$$

$$\begin{aligned} x_1 + 2x_2 &\le 6, \\ 2x_1 + x_2 &\le 8, \\ -x_1 + x_2 &\le 1, \\ x_2 &\le 2, \\ x_1 \ge 0, x_2 &\ge 0. \end{aligned}$$

Miền chấp nhận được được biểu diễn ở H.1.2, các số trong vòng tròn và mũi tên, chẳng hạn ① là chỉ số thứ tự của ràng buộc tương ứng và phía của nửa không gian xác định bởi ràng buộc.

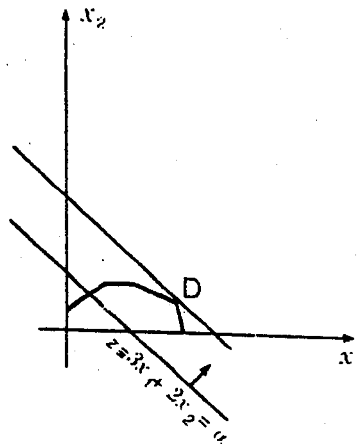


H. 1.2

Đường mức (level curve, trong trường hợp nhiều biến hơn thì gọi là *mặt mức* - level surface) của hàm mục tiêu là đường thẳng $z = 3x_1 + 2x_2 = \alpha$. Khi cho α tăng dần ta thấy điểm cuối cùng mà đường mức α còn cắt miền chấp nhận được là đỉnh D (H.1.3). D là giao điểm hai đường (1) và (2):

$$x_1 + 2x_2 = 6,$$
$$2x_1 + x_2 = 8.$$

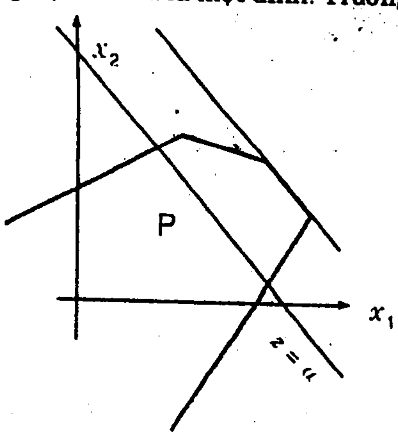
Giải hệ này ta được $x_1 = 3\frac{1}{3}$, $x_2 = 1\frac{1}{3}$, chính là *nghiệm chấp nhận được tối ưu* (optimal feasible solution), thường gọi tắt là *nghiệm tối ưu*. *Giá trị mục tiêu tối ưu* (optimal objective value) là $z^* = z(3\frac{1}{3}, 1\frac{1}{3}) = 12\frac{1}{3}$ ngàn USD.



H. 1.3

Qua bài toán cụ thể trên ta có thể nhận xét sơ bộ như sau:

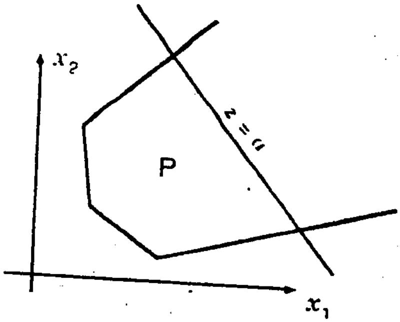
a. Miền chấp nhận được của quy hoạch tuyến tính là tập lồi đa diện. Nếu nó giới nội (tức là đa diện lồi) thì quy hoạch tuyến tính có nghiệm tối ưu là một đỉnh. Trường hợp nghiệm tối ưu không duy nhất nhưng miền chấp nhận được (thậm chí có thể không giới nội) có đỉnh (H.1.4) thì vẫn luôn có nghiệm tối ưu là đỉnh.



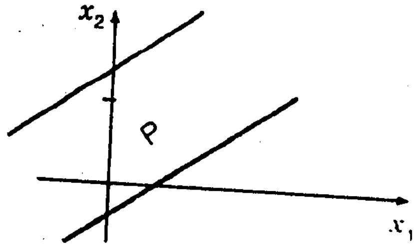
H. 1.4

b. Trường hợp không có nghiệm tối ưu thì hàm mục tiêu không giới nội trên miền chấp nhận được (H.1.5).

c. Trường hợp miền chấp nhận được không có đỉnh được minh họa ở (H.1.6). Bài toán quy hoạch tuyến tính có thể không có



H. 1.5



H. 1.6

nghiệm tối ưu hoặc có nhưng không có nghiệm tối ưu là đỉnh.

Những nhận xét hình học trên đây vẫn đúng cho cả trường hợp nhiều biến (hơn 2) và cho những gợi ý quan trọng để xây dựng phương pháp giải quy hoạch tuyến tính, ở dạng các thuật toán, bằng ngôn ngữ đại số để có thể lập trình cho máy tính.

Về nguyên tắc, ta có thể giải quy hoạch tuyến tính hai biến, với số ràng buộc bất kì, bằng hình học.

1.3.2 Bài toán với hai ràng buộc

(a) Giải bằng hình học

Đối lại theo một nghĩa nào đó với bài toán hai biến là bài toán với hai ràng buộc đẳng thức, còn số biến là bất kì. Người ta cũng tận dụng được yếu tố "hai" ở đây để đưa về biểu diễn và giải bài toán trong $\mathbf{R}^2$. Ta nêu phương pháp qua một thí dụ cụ thể.

Xét quy hoạch tuyến tính

$$\begin{aligned}
\min z = -12x_1 - 20x_2 - 18x_3 - 40x_4,\\
4x_1 + 9x_2 + 7x_3 + 10x_4 \le 6000,\\
x_1 + x_2 + 3x_3 + 40x_4 \le 4000,\\
x_1, x_2, x_3, x_4 \ge 0.
\end{aligned} \tag{1.6}$$

Trước hết, thêm biến bù x_5, x_6 để đưa bài toán về dạng chính tắc và viết lại như sau. Tìm cực tiểu z với ràng buộc $x_j \geq 0$, j = 1, ..., 6, và

$$
\begin{aligned}
4x_1 + 9x_2 + 7x_3 + 10x_4 + x_5 \quad &= 6000,\\
x_1 + x_2 + 3x_3 + 40x_4 \quad + x_6 &= 4000,\\
-12x_1 - 20x_2 - 18x_3 - 40x_4 \quad &= z.
\end{aligned}
$$

Để giải bằng hình học, ta biến đổi bài toán. Cộng hai ràng buộc đẳng thức thành một ràng buộc mới. Khi đó một bất kì trong hai ràng buộc cũ trở nên thừa (vì là tổ hợp tuyến tính của hai ràng buộc khác) và được bỏ đi. Chẳng hạn ta bỏ ràng buộc ứng với vế phải là 4000.

Tiếp theo, ta thay đổi "đơn vị" của các biến x_j, j = 1, ..., 6, như sau. Nhận xét rằng tổng hai hệ số trên cột một (của biến x_1) là 4 + 1 = 5. Tưởng tượng rằng nếu các biến khác đều bằng 0 thì $x_1 = 2000$ và $5x_1$ bằng tổng hai vế phải là 10000. Ta sẽ chọn biến mới thứ nhất y_1 sao cho khi các biến khác bằng 0 thì $y_1 = 1$, tức là $2000y_1 = x_1$. Cũng làm như vậy đối với các biến khác ta được sáu biến mới y_j, j = 1, ..., 6, quan hệ với các biến cũ như sau

$$
\begin{aligned}
&2000y_1 = x_1,\ 1000y_2 = x_2,\ 1000y_3 = x_3,\\
&200y_4 = x_4,\ 10000y_5 = x_5,\ 10000y_6 = x_6.
\end{aligned}
$$

Thực chất việc đổi biến này là thay đổi đơn vị đo (xem ý nghĩa của các biến ở các bài toán mở đầu).

Ta cũng đổi đơn vị vế phải ràng buộc và mục tiêu. 10000 đơn vị (thường là trữ lượng vật tư ... trong bài toán thực tế) thành một đơn vị mới và $10000\bar{z} = z$, ở đây $\bar{z}$ là giá trị theo đơn vị mới của mục tiêu. Khi đó bài toán trở thành: Tìm cực tiểu $\bar{z}$ với ràng buộc $y_j \geq 0$, j = 1, ..., 6 và

$$
\begin{aligned}
y_1 + \quad y_2 + \quad y_3 + \quad y_4 + y_5 + y_6 &= 1, \\
0{,}8y_1 + 0{,}9y_2 + 0{,}7y_3 + 0{,}2y_4 + y_5 \quad &= 0{,}6, \\
-2{,}4y_1 - \quad 2y_2 - 1{,}8y_3 - \quad 0{,}8y_4 \quad &= \bar{z}.
\end{aligned} \tag{1.7}
$$

Ký hiệu hệ số mỗi cột ở (1.7), trừ phần tử đầu tiên, là các điểm sau đây trong không gian hai chiều: $A_1 = \begin{pmatrix} 0{,}8 \\ -2{,}4 \end{pmatrix}$, $A_2 = \begin{pmatrix} 0{,}9 \\ -2 \end{pmatrix}$, $A_3 = \begin{pmatrix} 0{,}7 \\ -1{,}8 \end{pmatrix}$, $A_4 = \begin{pmatrix} 0{,}2 \\ -0{,}8 \end{pmatrix}$, $A_5 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $A_6 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $R = \begin{pmatrix} 0{,}6 \\ \bar{z} \end{pmatrix}$. Khi đó, về mặt hình học (1.7) có thể phát biểu tương đương là: Điểm R là tổ hợp lồi của $A_1, \ldots, A_6$ với các hệ số tổ hợp là $y_1, \ldots, y_6$. Ta thấy $\bar{z}$ là thay đổi theo $y_1, \ldots, y_6$ nhưng điểm R luôn nằm trên đường thẳng đứng (cũng ký hiệu là đường R) cắt trục hoành ở 0,6. Do đó, bài toán quy hoạch tuyến tính (1.6) có dạng hình học tương đương là: Tìm trên đường thẳng đứng cắt trục hoành ở 0,6 một điểm là tổ hợp lồi của $A_1, \ldots, A_6$ và có tung độ $\bar{z}$ thấp nhất. Từ H.1.7 ta thấy tập hợp tất cả các tổ hợp lồi có thể lập được từ các điểm $A_1, \ldots, A_6$, chính là đa giác có các đỉnh là $A_1, \ldots, A_6$ và điểm cần tìm là R^*. R^* nằm trên đoạn A_1A_4 nên chỉ là tổ hợp lồi của hai điểm A_1 và A_4, các hệ số khác phải bằng 0. Vậy ta có hệ phương trình để xác định y_1 và y_4 là

$$
\begin{aligned}
0{,}8y_1 + 0{,}2y_4 &= 0{,}6, \\
y_1 + \quad y_4 &= 1.
\end{aligned} \tag{1.8}
$$

Do đó quy hoạch tuyến tính (1.7) có nghiệm tối ưu là: $y_1 = \frac{2}{3}$, $y_4 = \frac{1}{3}$, $y_2 = 0$, $y_3 = 0$, $y_5 = 0$, $y_6 = 0$. Giá trị mục tiêu tối ưu tương ứng là $\bar{z} = -2{,}4y_1 - 0{,}8y_4 = -\frac{5{,}6}{3}$. Trở lại biến ban đầu, quy hoạch tuyến tính (1.6) có nghiệm tối ưu là $x_1 = 2000\,y_1 = \frac{4000}{3}$, $x_2 = 0$, $x_3 = 0$, $x_4 = 200\,y_4 = \frac{200}{3}$, $x_5 = 0$, $x_6 = 0$ và giá trị mục tiêu tối ưu là $z = 10000\bar{z} = -\frac{56000}{3}$.

(b) Giải thông qua bài toán đối ngẫu

Ta có thể giải bằng hình học bài toán với hai ràng buộc bằng cách suy luận mang tính đối ngẫu như sau. (Định nghĩa và nghiên cứu về đối ngẫu được trình bày ở Chương 6.)

Xét họ $\mathcal{L}$ các đường thẳng L trong mặt phẳng (u,v) sao cho mọi điểm $A_1, \ldots, A_6$ nằm ở L hoặc phía trên L. Phương trình của một đường L như vậy tổng quát có dạng

$$v = au + b.$$

Điểm $(\bar{u}, \bar{v})$ nằm ở đường L hoặc phía trên L nếu

$$\bar{v} \geq a\bar{v} + b.$$

Vì vậy mọi điểm $A_1, \ldots, A_6$ phải có tọa độ thỏa điều này, tức là a và b phải thỏa

$$\begin{aligned}
-2{,}4 &\geq 0{,}8a + b,\\
-2 &\geq 0{,}9a + b,\\
-1{,}8 &\geq 0{,}7a + b,\\
-0{,}8 &\geq 0{,}2a + b,\\
0 &\geq 1.a + b,\\
0 &\geq 0.a + b.
\end{aligned} \tag{1.9}$$

Giao điểm của L và R là điểm có dạng $S(0{,}6; \bar{v})$ với $\bar{v} = 0{,}6a + b$. Rõ ràng (xem H.1.7) giao điểm S như vậy với tung độ $\bar{v}$ lớn nhất có thể được chính là $\mathbf{R}^*$ và cho ta nghiệm tối ưu của quy hoạch tuyến tính. Vậy về hình học bài toán quy hoạch tuyến tính trở thành tìm giao điểm S của họ các đường L (trong $\mathcal{L}$) cắt R ở điểm cao nhất. Bài toán này có thể viết ở dạng giải tích như sau đây.

Tìm a, b để

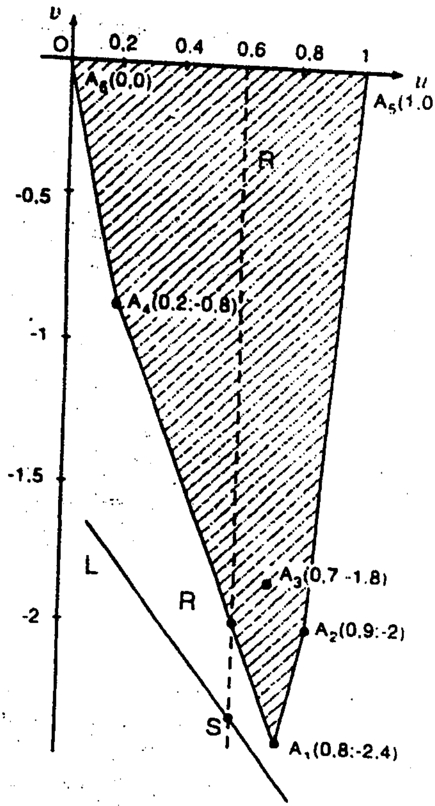
$$\begin{aligned}
&\max \bar{v} = 0{,}6a + b,\\
&0{,}8a + b \leq -2{,}4,\\
&0{,}9a + b \leq -2,
\end{aligned}$$

$$0{,}7a + b \le -1{,}8, \quad (1.10)$$
$$0{,}2a + b \le -0{,}8,$$
$$a + b \le 0,$$
$$b \le 0.$$

Đây lại là một bài toán quy hoạch tuyến tính, gọi là *bài toán đối ngẫu* (dual problem) của quy hoạch tuyến tính (1.7). (Hãy nhận xét so sánh dạng hai bài toán!) Bài toán (1.7) sẽ gọi là *bài toán gốc* (primal problem), hoặc quy hoạch tuyến tính gốc. Rõ ràng tung độ của $\mathbf{R}^*$ cho ta giá trị mục tiêu tối ưu chung của cả hai bài toán gốc và đối ngẫu, tức là ta có

$$\max \bar{v} = \min \bar{z}.$$

Ở Chương 6, ta sẽ thấy kết luận như vậy cho trường hợp hai bài toán đối ngẫu tổng quát chính là nội dung của Định lí đối ngẫu Von Neumann nổi tiếng.



H. 1.7

Cách giải hình học bằng suy luận đối ngẫu như trên có thể được thực hiện một cách giải tích như sau.

Giả sử từ hình học hoặc từ nhận xét nào đó ta phỏng đoán là một cặp điểm A_1, A_4 chẳng hạn sẽ sinh ra nghiệm tối ưu của quy hoạch tuyến tính như trên. Ta có thể kiểm tra điều này bằng cách kiểm tra hai điều kiện:

(i) Xét xem giao của đường nối A_1 và A_4 với đường thẳng đứng R có nằm giữa A_1 và A_4 không, bằng cách đặt y_2, y_3, y_5 và y_6 bằng 0 trong hai ràng buộc của quy hoạch tuyến tính gốc (1.7) ta được (1.8) và giải ra y_1 và y_4 xem có phải không âm hay không. (Trong thí dụ đang xét ta được $y_1 = \frac{2}{3}$, $y_4 = \frac{1}{3}$.)

(ii) Xét xem tất cả các điểm $A_1, ..., A_6$ có nằm ở đường nối A_1A_4 hoặc trên đường này không. Muốn vậy ta xác định hệ số a và b của đường thẳng v = au + b qua A_1 và A_4 bằng cách thay tọa độ A_1 và A_4 vào:

$$0{,}8a + b = -2{,}4$$
$$0{,}2a + b = -0{,}8$$

và giải ra được $a = -\frac{8}{3}$, $b = \frac{5{,}6}{3}$; rồi thay giá trị này vào (1.9), tức là ràng buộc của bài toán đối ngẫu (1.10), xem có thỏa không. (Trong thí dụ đang xét ta thấy $a = -\frac{8}{3}$, $b = \frac{5{,}6}{3}$ thỏa (1.9).)

Nếu hai điều kiện đều thỏa, ta kết luận $y_1 = \frac{2}{3}$, $y_4 = \frac{1}{3}$, $y_2 = 0$, $y_3 = 0$, $y_5 = 0$, $y_6 = 0$ là nghiệm tối ưu.

1.3.3 Phương pháp Fourier - Motzkin

(a) Thuật toán

Bây giờ ta trình bày một phương pháp có lẽ là cổ nhất, để giải quy hoạch tuyến tính tổng quát (không phải chỉ dùng cho trường hợp đặc biệt như ở hai phương pháp trên), do Fourier (Jean - Baptiste Joseph Fourier) đưa ra năm 1826, xem [Fourier 1826]. Nhưng mãi đến năm 1890 ông mới công bố rộng rãi, xem [Fourier 1890]. Độc lập với ông, T. S. Motzkin công bố phương pháp này trong [Motzkin 1936]. Nội dung của phương pháp là mỗi bước loại một biến cho đến khi không làm tiếp được ta sẽ được bài toán đơn giản hơn nhiều.

Ta hãy xét một thí dụ:

$$\begin{aligned} &\min z = -x_1 - 2x_2,\\ &2x_1 + x_2 \le 6,\\ &x_1 + x_2 \ge 2,\\ &-x_1 + x_2 \ge 3,\\ &x_1 \ge 0,\\ &x_2 \ge 0. \end{aligned} \tag{1.11}$$

Để khử biến x_2 ta để x_2 ở một vế và viết lại để các bất đẳng thức đều cùng chiều; đồng thời thay cho mục tiêu ta xét một bất đẳng thức tương ứng là $z \ge -x_1 - 2x_2$:

$$\begin{aligned} &6 - 2x_1 \ge x_2,\\ &x_2 \ge 2 - x_1,\\ &x_2 \ge 3 + x_1,\\ &x_1 \ge 0,\\ &x_2 \ge 0,\\ &x_2 \ge -\frac{x_1}{2} - \frac{z}{2}. \end{aligned} \tag{1.12}$$

Ghép mỗi bất đẳng thức ... $\ge x_2$ với từng bất đẳng thức dạng $x_2 \ge$ ..., ta được

$$\begin{aligned} &6 - 2x_1 \ge 2 - x_1,\\ &6 - 2x_1 \ge 3 + x_1,\\ &x_1 \ge 0,\\ &6 - 2x_1 \ge 0,\\ &6 - 2x_1 \ge -\frac{x_1}{2} - \frac{z}{2}. \end{aligned}$$

Như vậy là biến x_2 đã được khử. Bây giờ ta đơn giản hệ trên đây rồi lại khử x_1 như đã làm với x_2, hệ bất đẳng thức nhận được sẽ là

$$\begin{array}{lcccc}
4 \ge x_1, & & 4 \ge 0, & & 4 \ge 0, \\
1 \ge x_1, & & 1 \ge 0, & & 1 \ge 0, \\
x_1 \ge 0, & (1.13) \Rightarrow & 3 \ge 0, & \Rightarrow & 3 \ge 0, \\
3 \ge x_1, & & 4 + \dfrac{z}{3} \ge 0. & & z \ge -12\,. \\
4 + \dfrac{z}{3} \ge x_1. & & & &
\end{array}$$

Ba bất đẳng thức đầu luôn thỏa nên hệ tương đương với $z \ge -12$. Không thể làm như các bước trên để khử biến z được nữa. Quá trình khử kết thúc. Chuyển sang quá trình thế ngược để nhận được giá trị tương ứng của các biến đã khử. Cho z một giá trị bất kì thỏa $z \ge -12$ thay vào (1.13) ta được các giá trị chấp nhận được của x_1. Chẳng hạn lấy $z = -12$, giá trị nhỏ nhất làm hệ ban đầu (1.12) là chấp nhận được (tức là có nghiệm), ta được từ (1.13) $0 \le x_1 \le 0$; suy ra $x_1 = 0$. thay $z = -12$, $x_1 = 0$ vào (1.12) ta được $6 \le x_2 \le 6$, tức là $x_2 = 6$. Vậy nghiệm chấp nhận được của quy hoạch tuyến tính (1.11) và làm cực tiểu mục tiêu z là $x_1 = 0$, $x_2 = 6$ và giá trị mục tiêu tối ưu là $z = -12$.

Đối với các bài toán cỡ nhỏ, phương pháp Fourier - Motzkin khá tiện lợi. Nhược điểm lớn nhất của phương pháp này là sau mỗi bước, tuy giảm được một biến nhưng có thể số ràng buộc tăng thêm khá nhiều. Đó là vì để khử x_k chẳng hạn, ta phải ghép mỗi bất đẳng thức $\ldots \ge x_k$ với từng bất đẳng thức $x_k \ge \ldots$. Vì thế phương pháp Fourier - Motzkin không thể cạnh tranh được với phương pháp đơn hình của Dantzig được trình bày sau đây.

Từ thí dụ trên ta có thể phát biểu thuật toán tổng quát như sau, để tìm nghiệm chấp nhận được của hệ bất đẳng thức.

THUẬT TOÁN FOURIER - MOTZKIN

(1) Viết lại các ràng buộc (tức là các bất đẳng thức)

$$\sum_{j=1}^{n} a_{ij}x_j \geq b_i,\ i = 1, ..., m;$$

ở dạng

$$a_{in}x_n \geq -\sum_{j=1}^{n-1} a_{ij}x_j + b_i,\ i = 1, ..., m:$$

Nếu $a_{in} \neq 0$ thì chia hai vế cho a_{in}. Ký hiệu $\bar{x} = (x_1, ..., x_{n-1})$. Ta nhận được dạng tương đương của hệ ràng buộc (tức là chúng cùng xác định tập các điểm chấp nhận được là một tập lồi đa diện P trong $\mathbf{R}^n$):

$$x_n \geq d_i + f_i^T\bar{x} \qquad \text{nếu } a_{in} > 0, \tag{1.14}$$

$$d_j + f_j^T\bar{x} \geq x_n, \qquad \text{nếu } a_{jn} < 0, \tag{1.15}$$

$$0 \geq d_k + f_k^T\bar{x} \qquad \text{nếu } a_{kn} = 0. \tag{1.16}$$

ở đây d_i, d_j và d_k là các số thực, các f_i, f_j và f_k là các vectơ trong $\mathbf{R}^{n-1}$.

(2) Giả sử Q là tập lồi đa diện trong $\mathbf{R}^{n-1}$ xác định bởi hệ ràng buộc

$$d_j + f_j^T\bar{x} \geq d_i + f_i^T\bar{x} \qquad \text{nếu } a_{in} > 0 \text{ và } a_{jn} < 0, \tag{1.17}$$

$$0 \geq d_k + f_k^T\bar{x} \qquad \text{nếu } a_{kn} = 0. \tag{1.18}$$

Trên đây là một bước lặp của thuật toán (được viết cụ thể cho bước thứ nhất để khử x_n). Qua mỗi bước lặp một biến được khử.

Để áp dụng thuật toán Fourier - Motzkin vào việc giải quy hoạch tuyến tính ta đưa thêm biến z vào cùng với bất đẳng thức tương ứng với mục tiêu ở dạng $z \geq \sum_{j=1}^{n} c_jx_j$ rồi giải hệ n + 1 biến $x_1, ..., x_n$, z bằng thuật toán Fourier - Motzkin. Giả sử đến bước cuối cùng ta được tập lồi đa diện Q trong $\mathbf{R}^1$ là không gian một biến z. Khi đó giá trị mục tiêu tối ưu là số z nhỏ nhất của tập Q này. Bằng cách thế ngược ta xác định được $x_1, x_2, ..., x_n$ của nghiệm tối ưu tương ứng. (Hãy so sánh với thí dụ cụ thể ở trên!)

(b) Giải thích hình học

Ta sẽ thấy chìa khóa của phương pháp Fourier - Motzkin là khái niệm chiếu xuống không gian con được định nghĩa như sau. *Phép chiếu* (projection) π_k từ $\mathbf{R}^n$ vào $\mathbf{R}^k$, $k \le n$, là ánh xạ biến mỗi vectơ $x = (x_1, ..., x_n)$ thành vectơ $(x_1, ..., x_k) \in \mathbf{R}^k$ (là k tọa độ đầu của x). Cho $S \subset \mathbf{R}^n$, ta định nghĩa chiếu của tập S xuống $\mathbf{R}^k$ là

$$\Pi_k(S) := \{ \pi_k(x) : x \in S \}.$$

Dễ kiểm tra rằng

$$\Pi_k(S) = \{ (x_1, ..., x_k) \in \mathbf{R}^k : \exists x_{k+1}, ..., \exists x_n \text{ sao cho } (x_1, ..., x_k, x_{k+1}, ..., x_n) \in S \}.$$

Giả sử ta phải xem xét tập lồi đa diện $P \subset \mathbf{R}^n$ có khác trống hay không. Rõ ràng $P \neq \varnothing$ tương đương với $\Pi_{n-1}(P) \neq \varnothing$. Vì vậy áp dụng liên tiếp các phép chiếu lên không gian con và cuối cùng đi đến $\Pi_1(P) \subset \mathbf{R}^1$. Tập lồi đa diện trong $\mathbf{R}^1$ là một đoạn (hữu hạn hoặc vô hạn) nên kiểm tra $\Pi_1(P)$ có khác rỗng hay không rất dễ.

Bây giờ ta chứng minh rằng mỗi bước lặp của thuật toán Fourier - Motzkin là một phép chiếu xuống không gian con kém đi một chiều (ứng với biến được khử).

ĐỊNH LÍ 1.3.1. *Tập lồi đa diện Q xây dựng theo một bước lặp của thuật toán Fourier - Motzkin mô tả trên đây là hình chiếu $\Pi_{n-1}(P)$ của P.*

CHỨNG MINH. Giả sử $\bar{x} \in \Pi_{n-1}(P)$. Khi đó sẽ có x_n sao cho $(\bar{x}, x_n) \in P$. Vectơ $(\bar{x}, x_n)$ này thỏa (1.14), (1.15), (1.16). Do đó $\bar{x}$ thỏa (1.17), (1.18), tức là $\bar{x} \in Q$. Vậy $\Pi_{n-1} \subset Q$.

Ngược lại, giả sử $\bar{x} \in Q$. Từ (1.17) ta thấy

$$\min_{\{j:\, a_{jn} < 0\}} d_j + f_j^T \bar{x} \geq \max_{\{i:a_{in}>0\}} d_i + f_i^T \bar{x}$$

Gọi x_n là một số bất kì nằm giữa hai vế bất đẳng thức này. Khi đó $(\bar{x}, x_n)$ thỏa (1.14), (1.15) và (1.16), tức là $(\bar{x}, x_n) \in P$. Vậy $Q \subset \Pi_{n-1}(P)$.□

Nhận xét rằng với $x = (x_1, ..., x_n)$ bất kì thì

$$\pi_{n-2}(\pi_{n-1}(x)) = (x_1, ..., x_{n-2}) = \pi_{n-2}(x).$$

Do đó với tập lồi đa diện P ta cũng có

$$\Pi_{n-2}(\Pi_{n-1}(P)) = \Pi_{n-2}(P).$$

Tổng quát hóa nhận xét này ta thấy khi áp dụng k bước lặp thì từ P ta được $\Pi_{n-k}(P)$; và nếu áp dụng n – 1 lần ta được $\Pi_1(P)$ là tập lồi đa diện một chiều rất đơn giản. Tiếc rằng qua các bước lặp số ràng buộc có thể tăng mạnh. Người ta có thí dụ cho thấy số ràng buộc có thể tăng theo hàm mũ theo n. Vì $\Pi_1(P) \subset \mathbf{R}^1$ mà lại có thể được mô tả bằng rất nhiều ràng buộc, hầu hết các ràng buộc là thừa, nhưng việc xác định cái nào thừa cũng là phức tạp khi bài toán ban đầu là cỡ lớn.

Một nhận xét quan trọng nữa là $\Pi_k(P)$ là tập lồi đa diện với $k \leq n$, vì việc thực hiện thuật toán khử biến luôn biến tập lồi đa diện thành tập lồi đa diện. Vậy ta có

HỆ QUẢ 1.3.2. *Giả sử $M \subset \mathbf{R}^{n+k}$ là tập lồi đa diện. Khi đó tập $\Pi_k(M)$, tức là tập*

$$\{x \in \mathbf{R}^n: \text{tồn tại } y \in \mathbf{R}^k \text{ để } (x,y) \in M\}$$

cũng là tập lồi đa diện.

HỆ QUẢ 1.3.3. *Ảnh qua ánh xạ tuyến tính của tập lồi đa diện $P \subset \mathbf{R}^n$,*

tức là tập $\{Bx: x \in P\}$ với B là ma trận cấp m x n, cũng là tập lồi đa diện.

CHỨNG MINH. Ta thấy

$$\{Bx: x \in P\} = \{y \in \mathbf{R}^m: \text{tồn tại } x \in P, Bx = y\}$$

chính là chiếu của tập lồi đa diện $S = \{(x,y) \in \mathbf{R}^{n+m}: x = y, x \in P\}$ lên không gian con với thành phần tọa độ y, nên cũng là tập lồi đa diện. (S là tập lồi đa diện vì là giao của không gian con $\{(x,y) \in \mathbf{R}^{n+m}: x - y = 0\}$ và tập lồi đa diện $\{(x,y) \in \mathbf{R}^{n+m}: x \in P, y \text{ tùy ý}\}$. □

HỆ QUẢ 1.3.4. ***Bao lồi của hữu hạn điểm trong $\mathbf{R}^n$ là một tập lồi đa diện.***

CHỨNG MINH. Cho hữu hạn điểm $x^1, \ldots, x^k \in \mathbf{R}^n$. Bao lồi $\{\sum_{i=1}^{k}\gamma_i x^i : \sum_{i=1}^{k}\gamma_i = 1, \gamma_i \geq 0\}$ là ảnh của tập lồi đa diện $\{(\gamma_1, \ldots, \gamma_k) \in \mathbf{R}^k: \sum_{i=1}^{k}\gamma_i = 1, \gamma_i \geq 0\}$ qua ánh xạ tuyến tính $(\gamma_1, \ldots, \gamma_k) \to \sum_{i=1}^{k}\gamma_i x^i$ (từ $\mathbf{R}^k$ vào $\mathbf{R}^n$), vậy sẽ là tập lồi đa diện theo Hệ quả 1.3.3. □

CHÚ Ý. Thực ra thuật toán Fourier - Motzkin có thể bị dừng khi vẫn còn nhiều biến chưa được khử. Cụ thể, khi tiến hành thuật toán có hai khả năng xảy ra

(i) đến một bước nào đó ta không thể khử được biến tiếp theo x_k vì tất cả các ràng buộc đều cùng ở dạng $x_k \geq \ldots$, hoặc cùng ở dạng $\ldots \geq x_k$. Khi đó dễ dàng tìm được một nghiệm chấp nhận được;

(ii) tất cả các biến đều bị khử, tức là ta nhận được dạng $\sum_j 0x_j \geq d$. Khi đó nếu $d \leq 0$ thì biến x_1 có thể lấy tùy ý và thế ngược lên hệ ràng buộc ở các bước trên ta được nghiệm chấp nhận được. Nếu $d > 0$ thì hệ không có nghiệm chấp nhận được.

Chú ý rằng phương pháp Fourier - Motzkin là để giải hệ bất đẳng thức tuyến tính (cũng gọi là hệ bất phương trình tuyến tính), nhưng có thể vận dụng để giải quy hoạch tuyến tính như trên đã nói. Bài toán quy hoạch tuyến tính có hàm mục tiêu tuyến tính (và hệ ràng buộc là các bất đẳng thức tuyến tính) chỉ được nghiên cứu từ khi Dantzig đưa ra mô hình vào năm 1947.

1.4 TRƯỜNG HỢP KHÔNG CÓ NGHIỆM CHẤP NHẬN ĐƯỢC

Phương pháp Fourier - Motzkin cho phép ta chứng minh khá đơn giản định lí về hệ ràng buộc tuyến tính không có nghiệm chấp nhận được như sau.

Ta gọi *bất đẳng thức không chấp nhận được* (infeasibility inequality) là bất đẳng thức có dạng

$$0.x_1 + 0.x_2 + \ldots + 0.x_n \geq d \text{ với } d > 0.$$

ĐỊNH LÍ 1.4.1. Hệ bất đẳng thức tuyến tính

$$\sum_{j=1}^{n} a_{ij}x_j \geq b_i,\ i = 1, \ldots, m, \tag{1.19}$$

không có nghiệm chấp nhận được khi và chỉ khi tồn tại một tổ hợp tuyến tính không âm (tức là với các hệ số không âm) của các bất đẳng thức đó, là một bất đẳng thức không chấp nhận được.

CHỨNG MINH. Định lí khẳng định là hệ (1.19) là không chấp nhận được khi và chỉ khi tồn tại $y_k \geq 0$, $k = 1, \ldots, m$, sao cho

$$\sum_{k=1}^{m} y_k a_{kj} = 0,\ j = 1, \ldots, n \text{ và } \sum_{k=1}^{m} y_k b_k > 0. \tag{1.20}$$

Nếu hệ (1.19) là không chấp nhận được thì khi dùng thuật toán Fourier - Motzkin ta sẽ được trường hợp (ii) nói trong chú ý ở cuối Mục 1.3, tức là được một bất đẳng thức không chấp nhận được. Xem lại thuật toán này ta thấy bước lặp chính là ta đã lập một tổ hợp tuyến tính không âm của các bất đẳng thức của hệ ban đầu.

Ngược lại, nếu có (1.20) thì rõ ràng hệ ban đầu (1.19) là không chấp nhận được vì (1.20) là các bất đẳng thức hệ quả của (1.19). □

CHÚ Ý. (i) Ở dạng ma trận, Định lí 1.4.1 nói là hệ $Ax \geq b$ là không chấp nhận được khi và chỉ khi tồn tại vectơ $y \geq 0$ sao cho $y^T Ax \geq y^T b$ là bất đẳng thức không chấp nhận được, tức là $y^T A = 0$, $y^T b > 0$.

(ii) Trường hợp tất cả các dấu $\geq$ là dấu $=$ ta được hệ quả sau

HỆ QUẢ 1.4.2. *Nếu hệ phương trình tuyến tính của các biến không âm là không chấp nhận được thì tồn tại một tổ hợp tuyến tính, của các phương trình của hệ, là một phương trình không chấp nhận được của các biến không âm.*

BÀI TẬP CHƯƠNG 1

1.1. Một máy cán thép có thể sản xuất hai sản phẩm thép tấm và thép cuộn với công suất cho mỗi loại là (nếu chỉ sản xuất một sản phẩm): thép tấm: 200 tấn/giờ, thép cuộn: 140 tấn/giờ. Lợi nhuận bán sản phẩm là: thép tấm 25 USD/tấn, thép cuộn: 30 USD/tấn. Theo tiếp thị, một tuần chỉ tiêu thụ được tối đa 6000 tấn thép tấm và 4000 tấn thép cuộn. Biết rằng máy làm việc 40 giờ một tuần. Vấn đề đặt ra là cần sản xuất mỗi loại sản phẩm bao nhiêu trong một tuần để có lợi nhuận cao nhất. Hãy diễn đạt bài toán thành quy hoạch tuyến tính. Bạn có thể giải bài toán bằng nhận xét trực tiếp không?

1.2. "Bài toán xe đạp" đặt ra như sau. Có n người cùng phải đi quãng đường 10 dặm mà chỉ có một xe đạp một chỗ ngồi. Tốc độ đi bộ của người j là w_j và đi xe đạp là b_j, $j = 1, ..., n$. Làm sao để thời gian người cuối cùng đến đích là ngắn nhất.

(a) Giải bài toán với $n=3$, $w_1=4$, $w_2=w_3=2$, $b_1 = 16$, $b_2 = b_3 = 12$.

(b) Chứng minh rằng giá trị mục tiêu tối ưu của quy hoạch tuyến tính

$$\min t$$

$$t - x_j - x'_j - y_j - y'_j \geq 0, \; j = 1, \ldots, n,$$

$$t - \sum_{j=1}^{n} y_j - \sum_{j=1}^{n} y'_j \geq 0,$$

$$w_j x_j - w_j x'_j + b_j y_j - b_j y'_j = 10, \; j = 1, \ldots, n,$$

$$\sum_{j=1}^{n} b_j y_j - \sum_{j=1}^{n} b_j y'_j \leq 10,$$

$$x_j, x'_j, y_j, y'_j \geq 0, \; j = 1, \ldots, n.$$

là nhỏ hơn hoặc bằng giá trị mục tiêu tối ưu của bài toán xe đạp.

1.3. Giả sử Y là một biến ngẫu nhiên nhận một trong n giá trị đã biết a_1, a_2, ..., a_n. Biết rằng hoặc là Y có phân phối xác suất p cho bởi $P(Y = a_j) = p_j$ hoặc q cho bởi $P(Y = a_j) = q_j$. Tất nhiên $p_j \geq 0$, $\sum_{j=1}^{n} p_j = 1$ và q_j cũng có tính chất đó. Ta muốn đoán xem Y có phân phối p hay phân phối q. Cụ thể là với mỗi giá trị a_j ta sẽ khẳng định với xác suất x_j rằng phân phối là p và xác suất $1 - x_j$ rằng phân phối là q. Ta muốn xác định các xác suất x_j, $j = 1, \ldots, n$, sao cho xác suất của sự kiện là ta khẳng định phân phối là p nhưng thực tế nó là q, không lớn hơn β, ở đây β là số nhỏ (chẳng hạn 0,05) cho trước. Hơn nữa, ta muốn làm cực đại xác suất của sự kiện là ta nói phân phối là p và thực tế nó đúng là p. Hãy phát biểu bài toán này thành quy hoạch tuyến tính.

1.4. Một nhà máy chế biến thịt sản xuất ba loại thịt bò, lợn và cừu, với tổng lượng mỗi ngày 480 tấn bò, 400 tấn lợn và 230 tấn cừu. Mỗi loại đều có thể bán được ở dạng tươi hoặc nấu chín. Tổng lượng các loại thịt có thể nấu chín (để bán) trong giờ làm việc là 420 tấn. Ngoài ra, có thể nấu thêm ngoài giờ 250 tấn (với giá cao hơn). Lợi nhuận thu được trên một tấn được cho bằng bảng sau, với đơn vị là triệu đồng.

	tươi	nấu chín	nấu chín ngoài giờ
bò	8	14	11
lợn	4	12	7
cừu	4	13	9

Thí dụ: Phương án sản xuất sau đây sẽ cho lợi nhuận 9965 triệu đồng

	tươi	nấu chín	nấu chín ngoài giờ
bò	165 tấn	280 tấn	35 tấn
lợn	295 tấn	70 tấn	35 tấn
cừu	55 tấn	70 tấn	105 tấn

Mục đích của nhà máy là tìm phương án sản xuất để làm cực đại lợi nhuận. Hãy phát biểu bài toán ở dạng quy hoạch tuyến tính chuẩn.

1.5. Một xưởng mộc làm bàn và ghế. Để một công nhân làm xong một bàn cần hai giờ còn một ghế thì cần 30 phút. Khách hàng thường mua nhiều lắm là bốn ghế kèm theo một bàn. Do đó, tỉ lệ sản xuất ghế và bàn nhiều nhất là 4:1. Giá bán một bàn là 135 USD và một ghế là 50 USD. Hãy xác định xem số lượng sản xuất ghế và bàn hàng ngày là bao nhiêu để làm cực đại doanh thu. Biết rằng xưởng có bốn công nhân làm việc 8 giờ một ngày.

1.6. Một nhà máy sản xuất hai kiểu mũ. Thời gian lao động để làm ra một mũ kiểu thứ nhất nhiều gấp hai lần thời gian để làm xong mũ kiểu hai. Nếu sản xuất toàn mũ kiểu thứ hai thì nhà máy làm được 500 mũ một ngày. Thị trường tiêu thụ được trong mỗi ngày nhiều nhất là 150 mũ kiểu một và 200 mũ kiểu hai. Tiền lãi một mũ kiểu một là 8 USD và kiểu hai là 5 USD. Cần sản xuất bao nhiêu mũ mỗi kiểu trong ngày để tổng tiền lãi lớn nhất.

1.7. *Bài toán gà và trứng của Dantzig.* Trong hai tuần một con gà mái đẻ được 12 trứng để bán hoặc ấp được 4 trứng nở ra gà con. Sau 8 tuần thì bán tất cả gà và trứng với giá 0,6 USD một gà (lớn hoặc bé đều cùng giá) và 0,1 USD một trứng. Cần phải bố trí gà đẻ và ấp trứng như thế nào để doanh thu lớn nhất. Phát biểu bài toán cho hai trường hợp:

a) ban đầu có 100 gà mái và 100 trứng;

b) ban đầu có 100 gà mái và không có trứng.

1.8. Một công ty điện tử sản xuất hai kiểu radio trên hai dây chuyền độc lập. Công suất của dây chuyền một là 60 radio/ngày và dây chuyền hai là 75 radio/ngày. Để sản xuất một chiếc radio kiểu một cần 10 linh kiện điện tử E, và một chiếc radio kiểu hai cần 8 linh kiện này. Số linh kiện này được cung cấp mỗi ngày không quá 800. Tiền lãi khi bán một radio kiểu một là 30 USD và kiểu hai là 20 USD. Xác định phương án sản xuất cho lãi nhiều nhất trong ngày.

1.9. Hãy đưa bài toán

$$\min 2x_1 + |3x_2 - 10|,$$
$$|x_1 + 2| + |x_2| \le 5,$$

về bài toán quy hoạch tuyến tính tương đương.

1.10. Giải bằng hình học

(a)
$$\begin{aligned} &\min -x_1 + x_2,\\ &-x_1 - 2x_2 \le 6,\\ &x_1 - 2x_2 \le 4,\\ &-x_1 + x_2 \le 1,\\ &x_1, x_2 \le 0; \end{aligned}$$

(b)
$$\begin{aligned} &\max x_1 - x_2,\\ &3x_1 + x_2 \ge 3,\\ &x_1 + 2x_2 \ge 4,\\ &x_1 - x_2 \le 1,\\ &x_1 \le 5,\\ &x_2 \le 5. \end{aligned}$$

(c)
$$\begin{aligned} &\max 5x_1 + 6x_2,\\ &x_1 - 2x_2 \ge 2,\\ &-2x_1 + 3x_2 \ge 2, \end{aligned}$$

(x_1, x_2 không bị ràng buộc dấu);

(d)
$$\begin{aligned} \min -2x_1 - x_2,& \\ x_1 + 2x_2 &\le 6, \\ x_1 - x_2 &\ge 3, \\ x_1, x_2 &\ge 0. \end{aligned}$$

1.11. Chứng minh rằng bài toán tối ưu (với $x \in R^n$)
$$\begin{aligned} &\max f(x), \\ &Ax = b, \\ &x \ge 0, \end{aligned}$$
trong đó $f(x) = \min \{c^1x, c^2x, ..., c^p x\}$ có thể đưa về bài toán quy hoạch tuyến tính
$$\begin{aligned} &\max x_{n+1}, \\ &Ax = b, \\ &x \ge 0, \\ &x_{n+1} \le c^1x,\ x_{n+1} \le c^2x, ..., x_{n+1} \le c^px. \end{aligned}$$

1.12. Chỉ ra bằng cách biểu diễn hình học rằng quy hoạch tuyến tính
$$\begin{aligned} \max 3x_1 + 2x_2,& \\ 2x_1 + x_2 &\le 2, \\ 3x_1 + 4x_2 &\ge 12, \\ x_1, x_2 &\ge 0, \end{aligned}$$
không có đỉnh chấp nhận được. Có thể nói gì về nghiệm tối ưu của bài toán.

1.13. Giải bằng hình học và bằng phương pháp Fourier - Motzkin
$$\begin{aligned} \min 9x_1 + 8x_2,& \\ x_1 - 2x_2 &\le 3, \\ 3x_1 - 4x_2 &\ge 5, \\ 6x_1 - 7x_2 &= 8, \\ x_1, x_2 &\ge 0. \end{aligned}$$

1.14. Tìm điều kiện cần và đủ đối với các số thực a và b để quy hoạch tuyến tính

$$\max x_1 + x_2,$$
$$ax_1 + bx_2 \le 1;$$
$$x_1, x_2 \ge 0;$$

(i) có nghiệm chấp nhận được;

(ii) có nghiệm tối ưu;

(iii) không giới nội.

1.15. Tìm mối liên hệ giữa hai quy hoạch tuyến tính

$$\min c^Tx,\ Ax = b,\ x \ge 0,$$

và

$$\min \alpha c^Tx,\ Ax = \lambda b,\ x \ge 0,$$

ở đây $\alpha > 0$ và $\lambda > 0$ là hai số cho trước.

1.16. Giải bằng hình học

(a)
$$\begin{aligned} \min\ & 2x_1 + x_2 + 2x_3 + x_4 + 4x_5, \\ & 4x_1 + 2x_2 + 13x_3 + 3x_4 + x_5 = 17, \\ & x_1 + x_2 + 5x_3 + x_4 + x_5 = 7, \\ & x_1, \dots, x_5 \ge 0; \end{aligned}$$

(b)
$$\begin{aligned} \min\ & 2x_1 + 3x_2 + x_3 + 5x_4 + x_5, \\ & 4x_1 + 2x_2 + 3x_3 + x_4 + 4x_5 \le 50, \\ & 3x_1 + 7x_2 + x_3 + 3x_4 + 2x_5 \le 100, \\ & x_1, \dots, x_5 \ge 0. \end{aligned}$$

1.17. Giải bằng phương pháp Fourier - Motzkin

(a)
$$\begin{aligned} \max\ & x_1 + 3x_2 + 2x_3 \\ (1)\ & x_1 + x_2 + x_3 = 1, \\ & 7x_1 + 2x_2 + 3x_3 \le 20, \\ & x_1 + 5x_2 + 4x_3 \le 30, \\ & x_1, x_2, x_3 \ge 0; \end{aligned}$$

(b)
$$\begin{aligned} \max\ & 3x_1 + x_2, \\ & -x_1 + x_2 \le -1, \\ & -3x_1 - x_2 \le -1, \\ & 4x_1 + 2x_2 \le 1 \\ & 2x_1 \le 1, \\ & x_1, x_2 \ge 0. \end{aligned}$$

(c)
$$\begin{aligned} \min\ & 2x_1 - x_2 + 3x_3 + 7x_4 - 5x_5, \\ & x_1 + 2x_2 + x_3 + x_4 + 6x_5 \ge 10, \\ & 2x_1 + 3x_2 + 4x_3 + x_4 + 2x_5 \ge 4, \\ & 3x_1 + 2x_2 + 3x_4 + x_5 \le 8, \\ & x_1, x_2 \ge 0. \end{aligned}$$

(d)
$$\begin{aligned} \max\ & 3x_1 + x_2, \\ & -x_1 + x_2 \le -1, \\ & -3x_1 - x_2 \le -1, \\ & 4x_1 + 2x_2 \le 1, \\ & 2x_2 \le 1, \\ & x_1, x_2 \ge 0. \end{aligned}$$

(e)
$$\begin{aligned} \min\ & x_1 + x_2, \\ & x_1 + 2x_2 \ge 2, \\ & 3x_1 + 2x_2 \le 1, \\ & x_1 + x_2 \ge 1, \\ & x_1, x_2 \ge 0. \end{aligned}$$

1.18. Giải quy hoạch có hai ràng buộc bằng hình học và bằng đối ngẫu

(a)
$$\begin{aligned} \max\ & 6x_1 + 8x_2 + 5x_3 + 9x_4, \\ & 2x_1 + x_2 + x_3 + 3x_4 \le 5, \\ & x_1 + 3x_2 + x_3 + 2x_4 \le 3, \\ & x_1, x_2, x_3, x_4 \ge 0; \end{aligned}$$

(b) $$\max -2x_1 - 14x_2 - 36x_3,$$
$$\begin{aligned} -2x_1 + x_2 + 4x_3 - x_4 &= 5,\\ x_1 + 2x_2 + 3x_3 + 3x_4 - x_5 &= -2,\\ x_1, \ldots, x_5 &\geq 0; \end{aligned}$$

(c) $$\min x_1 - 4x_2 + x_3 + 13x_4 + 23x_5,$$
$$\begin{aligned} 2x_1 + x_2 + x_3 + 4x_4 - x_5 &= 6,\\ x_1 + 2x_2 + x_3 + x_4 - 4x_5 &= 3,\\ x_1, \ldots, x_5 &\geq 0. \end{aligned}$$

(d) $$\min -14x_1 - 18x_2 - 16x_3 - 10x_4,$$
$$\begin{aligned} 4{,}5x_1 + 8{,}5x_2 + 6x_3 + 20x_4 + x_5 &= 6000,\\ x_1 + x_2 + 4x_3 + 40x_4 + x_6 &= 4000,\\ x_1, \ldots, x_6 &\geq 0. \end{aligned}$$

1.19. Hãy đưa bài toán

$$\min \sum_{j=1}^{n} c_j |x_j|,$$
$$\sum_{j=1}^{n} a_{ij} x_j = b_i,\ i = 1, \ldots, m,$$

về quy hoạch tuyến tính.

1.20. Giải bằng hình học

$$\min -x_1 + 4x_2,$$
$$\begin{aligned} -3x_1 + x_2 &\leq 6,\\ x_1 + 2x_2 &\leq 4,\\ x_1, x_2 &\geq 0; \end{aligned}$$

Từ đó suy ra nghiệm tối ưu của bài toán khác nhận được từ bài toán trên khi thay ràng buộc $x_1, x_2 \geq 0$ bởi $x_1 \geq -1$, $x_2 \geq -3$.

CHƯƠNG 2

PHƯƠNG PHÁP ĐƠN HÌNH

2.1 MỘT THÍ DỤ

Phương pháp đơn hình được G. Dantzig đưa ra năm 1947 (xem [Dantzig 1949], [Dantzig 1963]) cùng lúc khi ông khai sinh quy hoạch tuyến tính. Trong khoảng bốn chục năm, nó là phương pháp thực sự hiệu quả duy nhất để giải quy hoạch tuyến tính trong thực tế (khi các bài toán thường là cỡ lớn). Trong thập niên cuối vừa qua chỉ có phương pháp điểm trong là cạnh tranh được với phương pháp đơn hình. Với các cách nhìn hiện đại, tư tưởng của phương pháp đơn hình rất đơn giản. Ở đây ta sẽ theo cách tiếp cận của Vacek Chvátal (xem [Chvátal 1983]) và Robert J. Vanderbei (xem [Vanderbei 1996]).

Xét quy hoạch tuyến tính

$$
\begin{aligned}
\min z &= -5x_1 - 4x_2 - 3x_3,\\
2x_1 + 3x_2 + x_3 &\le 5,\\
4x_1 + x_2 + 2x_3 &\le 11,\\
3x_1 + 4x_2 + 2x_3 &\le 8,\\
x_1, x_2, x_3 &\ge 0.
\end{aligned}
\tag{2.1}
$$

Trước hết ta hãy đưa bài toán về dạng chính tắc bằng cách đưa vào biến bù w_1, w_2, w_3 và viết hệ ở dạng

$$\begin{aligned}
\min z &= -5x_1 - 4x_2 - 3x_3,\\
w_1 &= 5 - 2x_1 - 3x_2 - x_3,\\
w_2 &= 11 - 4x_1 - x_2 - 2x_3,\\
w_3 &= 8 - 3x_1 - 4x_2 - 2x_3,\\
&x_1, x_2, x_3\ w_1, w_2, w_3 \geq 0.
\end{aligned}$$

Để bắt đầu ta cần có một nghiệm chấp nhận được. Dễ nhận thấy một nghiệm là

$$x_1 = 0,\ x_2 = 0,\ x_3 = 0,\ w_1 = 5,\ w_2 = 11,\ w_3 = 8.$$

Giá trị mục tiêu tương ứng là $z = 0$.

Trong bước lặp đầu tiên ta xem có thể cải tiến nghiệm xuất phát thế nào để được nghiệm chấp nhận được mới với mục tiêu z nhỏ hơn. Vì hệ số của x_1 ở hàm mục tiêu là âm, nếu ta tăng giá trị x_1 từ 0 lên giá trị dương càng lớn thì z càng giảm (nếu ta giữ $x_2 = 0$, $x_3 = 0$). Nhưng khi đó giá trị các biến bù cũng bị thay đổi. Sự thay đổi này không ảnh hưởng hàm mục tiêu nhưng x_1 chỉ được thay đổi sao cho $w_1, w_2, w_3 \geq 0$. Vì $x_2 = 0$, $x_3 = 0$ nên

$$\begin{aligned}
w_1 &= 5 - 2x_1 \geq 0, & & & x_1 &\leq \frac{5}{2},\\
w_2 &= 11 - 4x_1 \geq 0, & &\Rightarrow & x_1 &\leq \frac{11}{4},\\
w_3 &= 8 - 3x_1 \geq 0, & & & x_1 &\leq \frac{8}{3}.
\end{aligned}$$

Vậy ta lấy $x_1 = \dfrac{5}{2}$ và khi đó $w_1 = 0$. Ta được nghiệm chấp nhận được mới tốt hơn sau bước một là

$$x_1 = \frac{5}{2},\ x_2 = 0,\ x_3 = 0,\ w_1 = 0,\ w_2 = 1,\ w_3 = \frac{1}{2}$$

Giá trị mục tiêu là $z = -12{,}5$.

Nhận xét rằng bước lặp đầu dễ thực hiện vì trong nghiệm xuất phát có một nhóm biến bằng 0 và các biến khác biểu diễn qua chúng qua các ràng buộc. Do đó để vào bước 2 ta cũng viết lại bài toán ở dạng tương tự, tức là đổi vai trò của x_1 và w_1. Vậy x_1, w_2, w_3 và z sẽ biểu diễn qua w_1, x_2 và x_3. Muốn vậy, ở ràng buộc đầu ta giải x_1 ra theo w_1, x_2 và x_3

$$x_1 = \frac{5}{2} - \frac{1}{2}w_1 - \frac{3}{2}x_2 - \frac{1}{2}x_3. \qquad (2.2)$$

Tiếp theo ta thay x_1 trong biểu thức của w_2, w_3 và z bằng biểu thức (2.2) của nó. Cách tốt nhất để thực hiện việc thay này là làm cái gọi là *phép toán trên hàng* (row operation) đối với bài toán (2.1) như sau.

Đối với phương trình cho w_2 ta lấy nó trừ đi 2 lần phương trình cho w_1 rồi chuyển w_1 sang vế phải ta được

$$w_2 = 1 + 2w_1 + 5x_2.$$

Làm tương tự cho hàng w_3 và hàng z, ta sẽ viết lại được (2.1) ở dạng

$$\begin{array}{rl}
z = & -12{,}5 + 2{,}5w_1 + 3{,}5x_2 - 0{,}5x_3 \\
\hline
x_1 = & 2{,}5 - 0{,}5w_1 - 1{,}5x_2 - 0{,}5x_3 \\
w_2 = & 1 + 2w_1 + 5x_2 \\
w_3 = & 0{,}5 + 1{,}5w_1 + 0{,}5x_2 - 0{,}5x_3.
\end{array} \qquad (2.3)$$

Chú ý là ở đây ta thêm đường kẻ để phân biệt rõ hàm mục tiêu và không cần viết rõ ràng buộc về dấu nữa. Đồng thời từ cách viết này ta có được ngay nghiệm ở bước 1 bằng cách cho các biến "độc lập" giá trị 0 và đọc ra ngay được các giá trị tương ứng của các biến "phụ thuộc" $x_1 = 2{,}5$, $w_2 = 1$, $w_3 = 0{,}5$.

Bắt đầu bước lặp 2 ta lại nhận xét như bước 1. Bây giờ chỉ có hệ số của x_3 ở hàm mục tiêu là âm nên chỉ tăng x_3 từ 0 thành dương và giữ nguyên w_1 và x_2 bằng 0 là làm z giảm. (Nếu tăng x_2 chẳng hạn thì z tăng, tức là làm nghiệm hiện có "xấu đi".) Cũng như trước, ta tăng x_3 nhiều nhất có thể để w_2, x_1 và w_3 không âm. Từ phương trình w_2 ta thấy w_2 không bị ảnh hưởng theo x_3. Hai phương trình còn lại cho ta $x_3 \leq 5$ và $x_3 \leq 1$. Ta phải lấy $x_3 = 1$ và được nghiệm chấp nhận được mới

$$x_1 = 2;\ x_2 = 0,\ x_3 = 1,\ w_1 = 0,\ w_2 = 1,\ w_3 = 0$$

và giá trị mục tiêu tương ứng $z = -13$.

Để vào bước lặp 3 ta lại viết phương trình ở dạng x_1, w_2, x_3 và z biểu diễn theo w_1, x_2 và w_3. Thực hiện phép toán trên hàng đối với phương trình cho w_3 ta được:

$$x_3 = 1 + 3w_1 + x_2 - 2w_3.$$

Làm tương tự đối với các hàng khác ta được

$$\begin{aligned}
z &= -13 + w_1 + 3x_2 + w_3 \\
\hline
x_1 &= 2 - 2w_1 - 2x_2 + w_3 \\
w_2 &= 1 + 2w_1 + 5x_2 \\
x_3 &= 1 + 3w_1 + x_2 - 2w_3.
\end{aligned} \tag{2.4}$$

Bắt đầu bước 3, ta thấy không có hệ số nào ở hàm mục tiêu là âm nên ta không thể cho biến "độc lập" nào tăng từ 0 thành số dương (và giữ nguyên giá trị 0 cho biến "độc lập" khác) để giảm hàm mục tiêu. Phương pháp đơn hình phải dừng. Nhưng rất tốt là ta thấy ngay nghiệm chấp nhận được ở bước 2 là tối ưu. Thật vậy, hệ (2.4) là tương đương với quy hoạch tuyến tính ban đầu (2.1). Nhận xét (2.4) ta thấy, vì các biến phải không âm nên $z \geq -13$ với mọi nghiệm chấp nhận được của quy hoạch tuyến tính, mà nghiệm đã nhận được đã cho $z = -13$ nên nó là tối ưu. Vậy nghiệm tối ưu là $x_1 = 2$, $x_2 = 0$, $x_3 = 1$ và giá trị mục tiêu tối ưu là $z = -13$.

Các hệ phương trình dạng (2.3), (2.4) mà ta lập ở mỗi bước lặp gọi là các *từ vựng* (dictionary). Trừ z ra, các biến nằm ở vế trái các phương trình (tức là biến "phụ thuộc") ở mỗi bước lặp gọi là *biến cơ sở* ở bước đó (basic variable). Các biến ở vế phải, tức là biến "độc lập", được gọi là *biến không cơ sở* (nonbasic variable). Nghiệm nhận được khi cho các biến không cơ sở giá trị 0 được gọi là *nghiệm cơ sở* (basic solution). Vậy mỗi từ vựng xác định một nghiệm cơ sở tương ứng.

2.2 THUẬT TOÁN ĐƠN HÌNH TRÊN TỪ VỰNG

Xét quy hoạch tuyến tính ở dạng chuẩn

$$\min \sum_{j=1}^{n} c_j x_j$$

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i,\ i = 1, \dots, m, \tag{2.5}$$

$$x_j \ge 0,\ j = 1, \dots, n.$$

Việc đầu tiên là đưa biến bù vào và đặt tên mục tiêu là z:

$$z = \sum_{j=1}^{n} c_j x_j,$$

$$w_i = b_i - \sum_{j=1}^{n} a_{ij} x_j,\ i = 1, \dots, m. \tag{2.6}$$

Ta thấy ở thí dụ trên đây khi tiến hành thuật toán đơn hình, biến ban đầu và biến bù được xử lí như nhau, không phân biệt. Do đó ta ký hiệu lại thành một bộ biến x:

$$(x_1, \dots, x_n, w_1, \dots, w_m) := (x_1, \dots, x_n, x_{n+1}, \dots, x_{n+m}).$$

Khi đó (2.6) trở thành

$$z = \sum_{j=1}^{n} c_j x_j,$$

$$x_{n+i} = b_i - \sum_{j=1}^{n} a_{ij} x_j,\ i = 1, \ldots, m.$$

Đây là từ vựng xuất phát. Nội dung của thuật toán đơn hình là chuyển từ một từ vựng sang một từ vựng khác với giá trị mục tiêu tốt hơn. Mỗi từ vựng có m biến cơ sở và n biến không cơ sở. Hãy kí hiệu B là tập các chỉ số tương ứng với các biến cơ sở trong các chỉ số {1, 2, ..., n + m} và N là tập chỉ số của các biến không cơ sở. Ở từ vựng xuất phát thì N = {1, ..., n} và B = {n+1, ..., n+m}, nhưng chúng sẽ thay đổi sau mỗi bước. Ở mỗi bước, từ vựng đều có dạng

$$z = \bar{z} + \sum_{j \in N} \bar{c}_j x_j$$

$$x_i = \bar{b}_i - \sum_{j \in N} \bar{a}_{ij} x_j,\ i \in B.$$

Ở đây dấu gạch trên đầu ký tự để chỉ rằng đại lượng này thay đổi qua các bước.

Ở mỗi bước lặp đúng một biến từ không cơ sở trở thành biến cơ sở, được gọi là *biến vào* (entering variable), và đúng một biến cơ sở trở thành biến không cơ sở, gọi là *biến ra* (leaving variable). Biến vào được chọn trong các biến có hệ số mục tiêu (tức hệ số trong hàm mục tiêu) âm để làm giảm hàm mục tiêu. Nếu không có hệ số mục tiêu âm thì nghiệm nhận được ở bước lặp đó là tối ưu. Nếu có nhiều hệ số mục tiêu âm ta được phép lựa chọn. Một số quy tắc chọn biến vào được nghiên cứu ở Chương 3. Bây giờ ta chọn một cách tự nhiên (cũng là quy tắc thường dùng) là chọn biến có hệ số (âm) nhỏ nhất để hi vọng làm giảm hàm mục tiêu nhiều nhất. (Khi đó ta vẫn còn độ tự do khi có nhiều hệ số bằng nhau.)

Biến ra được chọn để đảm bảo tính không âm của các biến. Giả sử biến vào đã chọn là x_k, tức là giá trị của nó trở thành dương. Khi đó các biến đang là cơ sở sẽ bị thay đổi và bằng

$$x_i = \bar{b}_i - \bar{a}_{ik}x_k,\ i \in B.$$

x_k được phép lớn đến mức mọi $x_i \geq 0$, $i \in B$ tức là

$$\frac{1}{x_k} \geq \frac{\bar{a}_{ik}}{\bar{b}_i},\ i \in B$$

Rõ ràng x_k lớn nhất thỏa mọi bất đẳng thức này sẽ là

$$x_k = \left(\max_{i \in B} \frac{\bar{a}_{ik}}{\bar{b}_i}\right)^{-1}$$

Ở đây ta quy ước $\frac{0}{0} = 0$ và ta sẽ xét sau trường hợp có số $\bar{b}_i = 0$ và trường hợp không có tỉ số $\frac{\bar{a}_{ik}}{\bar{b}_i}$ nào dương. Vậy quy tắc chọn biến ra là chọn biến có chỉ số $l \in B$ mà $\frac{\bar{a}_{lk}}{\bar{b}_k} = \left(\max_{i \in B} \frac{\bar{a}_{ik}}{\bar{b}_i}\right)$ (Ta vẫn còn tự do khi có nhiều tỉ số cùng bằng giá trị cực đại này.)

Sau khi chọn biến vào và biến ra, việc chuyển từ vựng sang từ vựng mới là nhờ các phép toán hàng. Toàn bộ việc làm này gọi là *phép xoay* (pivot). Vì có thể có nhiều biến vào và biến ra có thể lấy đều đảm bảo giảm hàm mục tiêu và các biến vẫn không âm, ta sẽ thấy có các quy tắc cụ thể để tránh sự không xác định đó, được gọi là *quy tắc xoay* (pivot rule).

2.3 TÌM TỪ VỰNG XUẤT PHÁT

Ở mục trước ta đã trình bày thuật toán đơn hình, nhưng ta chỉ xét trường hợp vế phải các ràng buộc đều không âm; nhờ vậy từ vựng xuất phát là chấp nhận được. (Ta quy ước như vậy nếu nghiệm tương ứng của từ vựng, khi cho các biến không cơ sở bằng 0 và đọc ra ngay được các biến khác, là chấp nhận được.)

Nếu quy hoạch tuyến tính chuẩn (2.5) có các $b_i < 0$ thì ta phải tìm từ vựng xuất phát chấp nhận được bằng một bài toán bổ trợ. Người ta tìm *bài toán bổ trợ* (auxiliary problem) cũng là quy hoạch tuyến tính nhưng có 2 tính chất quan trọng sau đây

(i) dễ thấy ngay từ vựng xuất phát chấp nhận được của bài toán bổ trợ này;

(ii) từ vựng tối ưu của nó cho ngay một từ vựng chấp nhận được của quy hoạch tuyến tính gốc đang cần xét.

Bài toán bổ trợ đó là

$$\min x_0,$$

$$\sum_{j=1}^{n} a_{ij}x_j - x_0 \le b_i,\ i = 1, \dots, m,$$

$$x_j \ge 0,\ j = 0, 1, \dots, n.$$

Để có một nghiệm chấp nhận được chỉ việc đặt $x_j = 0$, $j = 1, \dots, n$ và lấy $x_0 > 0$ đủ lớn. Dễ thấy quy hoạch tuyến tính cần xét có nghiệm chấp nhận được khi và chỉ khi bài toán bổ trợ có nghiệm chấp nhận được với $x_0 = 0$. Nhưng đây rõ ràng là nghiệm tối ưu của bài toán bổ trợ. Vậy, quy hoạch tuyến tính cần xét có nghiệm chấp nhận được khi và chỉ khi bài toán bổ trợ có mục tiêu tối ưu bằng 0.

Mặc dù dễ thấy ngay một nghiệm chấp nhận được của bài toán bổ trợ như trên đã thấy, nhưng ta vẫn chưa có một từ vựng chấp nhận được của nó. Làm sao tìm? Ta hãy minh họa qua một thí dụ:

$$\min 2x_1 + x_2,$$

$$-x_1 + x_2 \le -1,$$

$$\begin{aligned} -x_1 - 2x_2 &\le -2, \\ x_2 &\le 1, \\ x_1, x_2 &\ge 0. \end{aligned}$$

Bài toán bổ trợ là

$$\begin{aligned} &\min x_0, \\ -x_1 + x_2 - x_0 &\le -1, \\ -x_1 - 2x_2 - x_0 &\le -2, \\ x_2 - x_0 &\le 1, \\ x_0, x_1, x_2 &\ge 0. \end{aligned}$$

Ta đưa vào biến bù và viết từ vựng xuất phát không chấp nhận được của bài toán bổ trợ (vì có vế phải ràng buộc âm):

$$\begin{array}{l} \zeta = \qquad\qquad\qquad x_0 \\ \hline w_1 = -1 + x_1 - x_2 + x_0 \\ w_2 = -2 + x_1 + 2x_2 + x_0 \\ w_3 = 1 \qquad - x_2 + x_0. \end{array}$$

Vì từ vựng này chưa chấp nhận được, ta chưa thực hiện bước lặp đầu bằng phép xoay của thuật toán đơn hình, mà cứ lấy x_0 là biến vào và chọn biến ra là biến w_2 "không chấp nhận được nhất":

$$\begin{array}{l} \zeta = 2 - x_1 - 2x_2 + w_2 \\ \hline w_1 = 1 \qquad + 3x_2 + w_2 \\ x_0 = 2 - x_1 - 2x_2 + w_2 \\ w_3 = 3 - x_1 - 3x_2 + w_2. \end{array}$$

Đây đã là từ vựng chấp nhận được nên ta bắt đầu thuật toán đơn hình, cụ thể phải đưa x_2 vào và w_1 ra khỏi cơ sở, sau khi xoay ta được

$$\zeta = \frac{4}{3} - x_1 + \frac{2}{3}w_1 + \frac{1}{3}w_2$$

$$x_2 = \frac{1}{3} \qquad - \frac{1}{3}w_1 + \frac{1}{3}w_2$$

$$x_0 = \frac{4}{3} - x_1 + \frac{2}{3}w_1 + \frac{1}{3}w_2$$

$$w_3 = 2 - x_1 + w_1.$$

Ở bước lặp tiếp theo ta đưa x_1 vào và x_0 ra khỏi cơ sở:

$$\zeta = 0 + x_0$$

$$x_2 = \frac{1}{3} \qquad - \frac{1}{3}w_1 + \frac{1}{3}w_2$$

$$x_1 = \frac{4}{3} - x_0 + \frac{2}{3}w_1 + \frac{1}{3}w_2$$

$$w_3 = \frac{2}{3} + x_0 + \frac{1}{3}w_1 - \frac{1}{3}w_2.$$

Vì không còn hệ số âm ở hàm mục tiêu, từ vựng này là tối ưu. Để được từ vựng chấp nhận được tương ứng của quy hoạch tuyến tính cần xét ban đầu, ta bỏ x_0 khỏi từ vựng này (vì giá trị mục tiêu tối ưu là $\zeta = x_0 = 0$) và tính giá trị tương ứng của mục tiêu z:

$$z = 2x_1 + x_2 = 3 + w_1 + w_2.$$

Vậy từ vựng chấp nhận được của quy hoạch tuyến tính cần xét là

$$\begin{aligned}
\zeta &= 3 + w_1 + w_2 \\
\hline
x_2 &= \frac{1}{3} - \frac{1}{3}w_1 + \frac{1}{3}w_2 \\
x_1 &= \frac{4}{3} + \frac{2}{3}w_1 + \frac{1}{3}w_2 \\
w_3 &= \frac{2}{3} + \frac{1}{3}w_1 - \frac{1}{3}w_2.
\end{aligned} \tag{2.7}$$

May mắn, một cách vô tình, ở thí dụ này là từ vựng này đã tối ưu, vì các hệ số mục tiêu đều dương. Nói chung thì ta chỉ được từ vựng xuất phát chấp nhận được.

Thông thường người ta ghép luôn việc giải bài toán bổ trợ vào thuật toán đơn hình và gọi là *pha I* (phase I). Quá trình trực tiếp áp dụng thuật toán đơn hình từ từ vựng xuất phát chấp nhận được nhận từ pha I gọi là *pha II* (phase II). Toàn bộ thuật toán gọi là *thuật toán hai pha* (two - phase algorithm).

2.4 TRƯỜNG HỢP MỤC TIÊU KHÔNG GIỚI NỘI

Tất nhiên có quy hoạch tuyến tính mà hàm mục tiêu có thể dần đến $-\infty$ trong miền chấp nhận được. Trường hợp này sẽ không có nghiệm tối ưu.

Làm sao phát hiện ra là quy hoạch tuyến tính đang xét có hàm mục tiêu không giới nội như vậy? Câu trả lời là: có thể phát hiện được, và ngay trong quá trình thực hiện thuật toán đơn hình.

Ta hãy xét kĩ mấy trường hợp "bất thường" mà khi xây dựng thuật toán đơn hình ở Mục 2.2 cho trường hợp "bình thường" thuận lợi ta đã phải gác lại. Ở đây, ta xét trường hợp không có tỉ số $\frac{\bar{a}_{ik}}{\bar{b}_i}$,

$i \in B$, nào dương. (Trường hợp có $\bar{b}_i = 0$ với $i \in B$ sẽ được xét ở Chương 3 dưới đây.) Tỉ số này gặp phải khi tìm biến ra sau khi đã xác định biến x_k là biến vào, tức là tăng từ 0 lên một số dương. Nhớ rằng lúc đó các biến cơ sở là

$$x_i = \bar{b}_i - \bar{a}_{ik}x_k \quad i \in B.$$

Mà $\bar{b}_i$ và $-a_{ik}$ là cùng dấu (vì $\frac{\bar{a}_{ik}}{\bar{b}_i} \leq 0$) và là không âm. Do đó mọi biến cơ sở x_i không thể từ không âm trở thành âm. Vậy biến vào có thể lấy giá trị lớn tùy ý để hàm mục tiêu tiến tới $-\infty$. Lúc này ta nói là hàm mục tiêu không giới nội (dưới), hoặc bài toán không giới nội.

Chẳng hạn, xét từ vựng

$$\begin{aligned}
z &= 7 - x_2 + x_1 \\
x_3 &= 2 + 2x_2 - x_1 \\
x_4 &= 6 \qquad\; + 3x_1 \\
x_5 &= \qquad\qquad 2x_1.
\end{aligned} \tag{2.8}$$

Biến vào là x_2 và các tỉ số $\frac{\bar{a}_{ik}}{\bar{b}_i}$ là $-\frac{2}{2}$, $-\frac{0}{6}$ và $\frac{0}{0}$. Vì không có tỉ số nào dương nên bài toán là không giới nội.

2.5 SỰ DUY NHẤT NGHIỆM TỐI ƯU

Khi đã biết là có tồn tại nghiệm tối ưu của một quy hoạch tuyến tính đang xét, chẳng hạn khi đã có một từ vựng tối ưu (tức là từ vựng cho ta nghiệm tối ưu), thì vấn đề nghiệm tối ưu có duy nhất không cũng là câu hỏi quan trọng hay được đặt ra trong thực tế.

Thuật toán đơn hình có một ưu việt là cho phép ta xét được tính duy nhất này một cách đơn giản.

Chẳng hạn, xét từ vựng tối ưu (2.7). Phương trình z (tức là phương trình đầu) cho thấy mọi nghiệm chấp nhận được có giá trị mục tiêu z = 3, tức là tối ưu, đều phải thỏa $w_1 = 0$, $w_2 = 0$. Ba phương trình sau cho thấy nếu $w_1 = 0$, $w_2 = 0$ thì bắt buộc $x_2 = \frac{1}{3}$, $x_1 = \frac{4}{3}$ và $w_3 = \frac{2}{3}$. Vậy nghiệm tối ưu là xác định duy nhất.

Ở mục 1.3.1 ta đã thấy bằng hình học là quy hoạch tuyến tính có thể có vô số nghiệm tối ưu. Thuật toán đơn hình cho phép ta mô tả tất cả các nghiệm tối ưu khá đơn giản như sau.

THÍ DỤ 2.5.1 Giả sử (bằng thuật toán đơn hình) ta nhận được từ vựng tối ưu sau

$$\begin{array}{l}
z = 5 \qquad\qquad\qquad\quad + x_3 \\
\hline
w_1 = 3 + x_2 - 2w_2 + 7x_3 \\
x_1 = 1 - 5x_2 + 6w_2 - 8x_3 \\
w_3 = 4 + 9x_2 + 2w_2 - x_3.
\end{array}$$

Dòng đầu tiên cho thấy mỗi nghiệm tối ưu phải thỏa điều kiện $x_3 = 0$ (nhưng không cần $x_2 = 0$ hoặc $w_2 = 0$). Với nghiệm như vậy, các dòng dưới của từ vựng cho ta

$$\begin{aligned}
w_1 &= 3 + x_2 - 2w_2,\\
x_1 &= 1 - 5x_2 + 6w_2,\\
w_3 &= 4 + 9x_2 + 2w_2.
\end{aligned}$$

Do đó, để nhận được nghiệm tối ưu ta có thể cho x_2 và w_2 giá trị không âm tùy ý miễn là đảm bảo $w_1 \geq 0$, $x_1 \geq 0$ và $w_3 \geq 0$, tức là x_2 và w_2 chỉ cần thỏa

$$\begin{aligned}
-x_2 + 2w_2 &\leq 3,\\
5x_2 - 6w_2 &\leq 1,
\end{aligned}$$

$$-9x_2 - 2w_2 \leq 4$$
$$x_2, w_2 \geq 0.$$

(Thực ra bất đẳng thức $-9x_2 - 2w_2 \leq 4$ là thừa vì nó luôn đúng với $x_2 \geq 0$, $w_2 \geq 0$.)

2.6 DẠNG BẢNG CỦA PHƯƠNG PHÁP ĐƠN HÌNH

Trên đây ta đã dùng từ vựng để ghi kết quả ở mỗi bước lặp của thuật toán đơn hình. Một cách khác, thậm chí phổ biến hơn là dùng bảng biểu diễn như sau. Để tiện lợi ta xét lại thí dụ cũ là quy hoạch tuyến tính (2.1). Ta viết lại dạng chính tắc của bài toán như sau (mục tiêu được ghi ở dạng một phương trình)

$$\begin{array}{rcl}
-z - 5x_1 - 4x_2 - 3x_3 & = & 0 \\
\hline
2x_1 + 3x_2 + x_3 + w_1 & = & 5 \\
4x_1 + x_2 + 2x_3 + w_2 & = & 11 \\
3x_1 + 4x_2 + 2x_3 + w_3 & = & 8.
\end{array} \qquad (2.9)$$

Với quy ước kết quả ở mỗi bước đều viết ở dạng như vậy, ta chỉ cần ghi các hệ số là đủ. Ta sẽ được bảng sau (hoàn toàn xác định bài toán 2.1)

x_1	x_2	x_3	w_1	w_2	w_3	
5	4	3	0	0	0	0
2	3	1	1	0	0	5
4	1	2	0	1	0	11
3	4	2	0	0	1	8

(2.10)

Theo cách viết bảng như vậy thì từ vựng (2.3) đạt được sau bước 1 sẽ là

$$
\begin{array}{cccccc|c}
x_1 & x_2 & x_3 & w_1 & w_2 & w_3 & \\
\hline
0 & \frac{7}{2} & -\frac{1}{2} & \frac{5}{2} & 0 & 0 & \frac{25}{2} \\
\hline
1 & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} & 0 & 0 & \frac{5}{2} \\
0 & -5 & 0 & -2 & 1 & 0 & 1 \\
0 & -\frac{1}{2} & \frac{1}{2} & -\frac{3}{2} & 0 & 1 & \frac{1}{2}
\end{array}
\qquad (2.11)
$$

Ta hãy xem bảng (2.11) nhận được từ bảng (2.10) như thế nào. Nói cách khác ta hãy xem bước 1 tiến hành trên bảng (2.10) thế nào để nhận được bảng (2.11).

Để chọn biến vào ta lấy số âm lớn nhất ở hàng đầu không kể số tận cùng bên phải (là giá trị của $-z$). Trong thí dụ này là -5, và x_1 là biến vào. Khi đó cột đầu được gọi là *cột xoay* (pivot column). Chú ý là nếu không có số âm ở hàng đầu thì ta kết luận bảng hiện tại là tối ưu và viết được ra ngay nghiệm tối ưu tương ứng của quy hoạch tuyến tính. Thuật toán kết thúc.

Tiếp theo là tìm biến ra. Theo lí luận bằng từ vựng ở Mục 2.1 ta phải xét các hàng i mà phần tử a_{ik} của nó trên cột xoay k là dương. Nếu mọi phần tử trên cột xoay là âm, thì theo Mục 2.4, quy hoạch tuyến tính không có nghiệm tối ưu vì hàm mục tiêu không giới nội dưới trên miền chấp nhận được. Ngược lại, ta phải chọn hàng $i = l$ mà $\frac{a_{lk}}{b_l}$ là lớn nhất, tức là tỉ số giữa phần tử trên cột xoay và phần tử cùng hàng nhưng ở cuối hàng là lớn nhất. Trong thí dụ của ta là hàng 1 (không kể hàng mục tiêu trên cùng):

x_1	x_2	x_3	w_1	w_2	w_3	
– 5	– 4	– 3	0	0	0	0
2	3	1	1	0	0	5
4	1	2	0	1	0	11
3	4	2	0	0	1	8

Hàng l được chọn ra được gọi là *hàng xoay* (pivot row) và phần tử ở giao của cột xoay và hàng xoay được gọi là *phần tử trục* (pivot).

Tiếp đến, như đã thấy ở Mục 2.1 ta phải thực hiện phép toán trên hàng. Thực hiện điều này trên bảng sẽ là chia hàng xoay cho phần tử trục để biến phần tử trục thành 1; rồi lấy mỗi hàng khác trừ đi hàng xoay mới, nhân với số trên cột xoay của hàng đó để biến cột xoay thành cột các số 0 ngoại trừ phần tử trục. Đến đây ta được bảng (2.11) và kết thúc một bước lặp.

Dạng bảng viết có gọn hơn dạng từ vựng nhưng không tường minh bằng, nên sau đây ta vẫn sử dụng từ vựng.

Các thuật ngữ "phép xoay, cột xoay, hàng xoay" có thể hiểu là liên quan đến việc đưa một biến vào và lấy một biến ra ở mỗi bước lặp tương tự như ta "xoay" từ vựng hoặc bảng đơn hình đang có ở bước đó. Thực ra các thuật ngữ này, cùng với chính tên của thuật toán là "đơn hình" có hình ảnh hình học rõ ràng gắn liền với phép xoay thực sự mà ta sẽ nói đến ở Mục 4.6.

BÀI TẬP CHƯƠNG 2

2.1. Giải Bài tập 1.10 bằng phương pháp đơn hình, viết ở dạng từ vựng và ở dạng bảng.

2.2. Giải Bài tập 1.13 và 1.16 bằng phương pháp đơn hình ở dạng từ vựng và dạng bảng.

2.3. Giải Bài tập 1.17 bằng phương pháp đơn hình.

2.4. Giải Bài tập 1.18 bằng phương pháp đơn hình.

2.5. Giải bằng phương pháp đơn hình

(a)
$$\begin{aligned}
&\min -2x_1 - x_2,\\
&2x_1 + x_2 \le 4,\\
&2x_1 + 3x_2 \le 3\\
&4x_1 + x_2 \le 5\\
&x_1 + 5x_2 \le 1,\\
&x_1, x_2 \ge 0;
\end{aligned}$$

(b)
$$\begin{aligned}
&\max 3x_1 + 2x_2,\\
&-x_1 + 2x_2 \ge -1,\\
&x_1 - x_2 \le 2,\\
&-2x_1 - x_2 \ge -6,\\
&x_1 \le 5,\\
&2x_1 + x_2 \le 16,\\
&x_1 + x_2 \le 12,\\
&x_1 + 2x_2 \le 21\\
&x_2 \le 10\\
&x_1, x_2 \ge 0,
\end{aligned}$$

(c)
$$\begin{aligned}
&\min x_1 + 2x_3 + x_5,\\
&x_1 + x_2 + x_3 + x_4 + x_5 = 5,\\
&x_2 + x_3 + x_4 - x_5 = 2,\\
&x_3 - x_4 + x_5 = 1,\\
&x_1, \dots, x_5 \ge 0;
\end{aligned}$$

(d)
$$\begin{aligned} &\min 5x_1 - 10x_2,\\ &-2x_1 + x_2 \le 1\\ &x_1 - x_2 \ge -2,\\ &3x_1 + x_2 \le 8,\\ &-2x_1 + 3x_2 \ge -9,\\ &4x_1 + 3x_2 \ge 0,\\ &x_1 \ge 0. \end{aligned}$$

2.6. Giải bằng thuật toán hai pha.

(a)
$$\begin{aligned} &\max 2x_1 - 6x_2,\\ &-x_1 - x_2 - x_3 \le -2,\\ &2x_1 - x_2 + x_3 \le 1,\\ &x_1, x_2, x_3 \ge 0; \end{aligned}$$

(b)
$$\begin{aligned} &\min -x_1 - 3x_2,\\ &x_1 + x_2 \ge 3,\\ &-x_1 + x_2 \le -1,\\ &x_1 + 2x_2 \le 4,\\ &x_1, x_2 \ge 0; \end{aligned}$$

(c)
$$\begin{aligned} &\max x_1 + 3x_2,\\ &-x_1 - x_2 \le -3,\\ &-x_1 + x_2 \le -1,\\ &-x_1 + 2x_2 \le 2,\\ &x_1, x_2 \ge 0. \end{aligned}$$

2.7. Biểu diễn hình học miền chấp nhận được của các Bài tập 2.5(a), (b), (d) và 2.6(b), (c) và chỉ ra các đỉnh tìm được qua các bước của thuật toán đơn hình đã thực hiện.

2.8. Tìm thí dụ cho thấy một biến vào ở bước lặp nào đó có thể lại bị ra khỏi cơ sở ở bước ngay sau đó. Chứng minh rằng một biến ra ở một bước lặp thì không thể là biến vào ở bước lặp tiếp theo.

2.9. Giải quy hoạch tuyến tính

(a)
$$\max -x_1 - 3x_2 - x_3,$$
$$\begin{aligned} 2x_1 - 5x_2 + x_3 &\le -5,\\ 2x_1 - x_2 + 2x_3 &\le 4,\\ x_1,\ x_2,\ x_3 &\ge 0; \end{aligned}$$

(b)
$$\min x_{12} + 8x_{13} + 9x_{14} + 2x_{23} + 7x_{24} + 3x_{34},$$
$$\begin{aligned} x_{12} + x_{13} + x_{14} &\ge 1,\\ -x_{12} + x_{23} + x_{24} &= 0,\\ -x_{13} - x_{23} + x_{34} &= 0,\\ x_{14} + x_{24} + x_{34} &\le 1,\\ x_{12}, x_{13}, \ldots, x_{34} &\ge 0; \end{aligned}$$

(c)
$$\min x_1 - x_2 - x_3 - x_4,$$
$$\begin{aligned} x_1 + x_2 - x_4 &\le 1,\\ -x_1 + x_2 + x_4 &\le 1,\\ x_2 + x_3 &= 1,\\ x_1 \ge 0;\ x_2 &\ge 0. \end{aligned}$$

(d)
$$\min 2x_1 + 3x_2 + 4x_3,$$
$$\begin{aligned} -2x_2 - 3x_3 &\ge -5,\\ x_1 + x_2 + 2x_3 &\le 4,\\ x_1 + 2x_2 + 3x_3 &\le 7,\\ x_1, x_2, x_3 &\ge 0; \end{aligned}$$

(e)
$$\max x_1 + 3x_2,$$
$$\begin{aligned} -x_1 - x_2 &\le -3\\ -x_1 + x_2 &\le -1\\ x_1 + 2x_2 &\le 2,\\ x_1, x_2 &\ge 0, \end{aligned}$$

2.10. Giải quy hoạch tuyến tính

$$\max \sum_{j=1}^{n} p_j x_j,$$

$$\sum_{j=1}^{n} q_j x_j \le \beta,$$

$$0 \le x_j \le 1,\ j = 1, \ldots, n.$$

Ở đây các số p_j, $j = 1,\ldots, n$ dương và có tổng bằng 1. Các số q_j cũng có tính chất đó. Hơn nữa, giả sử rằng (gần như không mất tính tổng quát)

$$\frac{p_1}{q_1} < \frac{p_2}{q_2} < \ldots < \frac{p_n}{q_n}.$$

Cuối cùng, biết rằng $\beta > 0$ và nhỏ. (Xem Bài tập 1.3 để thấy nguồn gốc của bài toán này.)

2.11. *Điều kiện cần và đủ để nghiệm tối ưu duy nhất.* Giả sử x^* là nghiệm tối ưu của quy hoạch tuyến tính dạng chính tắc. Giả sử B và N là tập chỉ số cơ sở và không cơ sở trong một cơ sở tối ưu tương ứng với x^*. Gọi I là tập các chỉ số không cơ sở mà $\bar{c}_j = 0$.

(a) Chứng minh rằng nếu $I = \varnothing$ thì x^* là nghiệm tối ưu duy nhất.

(b) Chứng minh rằng x^* là nghiệm tối ưu duy nhất khi và chỉ khi quy hoạch tuyến tính có giá trị mục tiêu tối ưu bằng 0:

$$\max \sum_{j \in I} x_j\ ,$$

$$Ax = b,$$

$$x_j = 0,\ j \in N \setminus I,$$

$$x_j \ge 0,\ j \in B \cup I.$$

2.12. Chứng minh kết luận sau đây là đúng hoặc chứng minh nó là sai: từ vựng chấp nhận được có hàng mục tiêu

$$z = \bar{z} + \sum_j \bar{c}_j x_j$$

là từ vựng tối ưu khi và chỉ khi $\bar{c}_j \le 0$ với mọi j (đối với bài toán max).

2.13. Chứng minh rằng nếu quy hoạch tuyến tính có hai nghiệm tối ưu x^1 và x^2 thì mọi tổ hợp lồi

$$\lambda x^1 + (1 - \lambda)x^2, \ 0 \le \lambda \le 1,$$

đều là nghiệm tối ưu.

2.14. Chứng minh hoặc chỉ rằng mệnh đề sau là sai: Nếu bài toán ở dạng chuẩn

$$\max \sum_{j=1}^{n} c_j x_j, \quad Ax \le b, \quad x \ge 0,$$

là không giới nội, thì có một chỉ số k để bài toán $\max x_k$, $Ax \le b$, $x \ge 0$ cũng không giới nội.

2.15. Chứng minh rằng quy hoạch tuyến tính

$$\begin{aligned} &\max x_1 - x_2,\\ &-2x_1 + x_2 \le -1,\\ &-x_1 - 2x_2 \le -2,\\ &x_1, \quad x_2 \ge 0, \end{aligned}$$

là không giới nội, sau đó biểu diễn hình học.

CHƯƠNG 3

HIỆN TƯỢNG SUY BIẾN

Trong chương này ta sẽ xét trường hợp "bất thường" còn lại chưa giải quyết, nhưng đã nhắc đến, là khi chọn biến ra ta gặp trường hợp có tỉ số $\frac{\bar{a}_{ik}}{\bar{b}_i}$ không xác định vì $\bar{b}_i = 0$.

3.1 ĐỊNH NGHĨA VÀ THÍ DỤ

Bây giờ ta định nghĩa hiện tượng suy biến cho trường hợp ta dùng từ vựng khi tiến hành thuật toán đơn hình. Sau này, ở Mục 3.3 ta sẽ định nghĩa sự suy biến không thông qua từ vựng.

Ta nói một từ vựng là *suy biến* (degenerate) nếu có $i \in \overline{B}$ (tập chỉ số của các biến cơ sở) với $b_i = 0$. Một từ vựng suy biến có thể gây khó khăn khi ta thực hiện tiếp thuật toán đơn hình trên nó, nhưng cũng có thể không. Chẳng hạn, nếu ta nhìn lại từ vựng (2.8)

$$\begin{array}{l}
z = 7 - x_2 + x_1 \\
\hline
x_3 = 2 + 2x_2 - x_1 \\
x_4 = 6 \qquad\quad + 3x_1 \\
x_5 = \qquad\qquad\ \ 2x_1.
\end{array}$$

Đây là từ vựng suy biến vì $\overline{B} = \{3,4,5\}$ và $\bar{b}_5 = 0$. Nhưng như ta đã biết ở

Mục 2.4, vì trong các tỉ số $\frac{\bar{a}_{ik}}{\bar{b}_i}$, là $-\frac{2}{2}$, $-\frac{0}{6}$ và $\frac{0}{0}$, không có tỉ số nào dương nên bài toán là không giới nội. Nếu bây giờ ta gặp bài toán khác có từ vựng hoàn toàn như vậy, chỉ khác là $2x_2$ ở phương trình của x_3 lại là $-2x_2$, thì theo quy tắc chọn biến ra, x_2 là biến ra. Tiếp theo, với phép toán trên hàng để "xoay" ta được từ vựng mới

$$\begin{array}{llr}
z = 6 + \frac{x_3}{2} + 2x_1 & & \\
\hline
x_2 = 1 & + 3x_1 & \\
x_4 = 6 & + 3x_1 & \\
x_5 = & 2x_1. &
\end{array}$$

Đây vẫn là một từ vựng suy biến, nhưng đã là tối ưu, nên ta được ngay nghiệm tối ưu của quy hoạch tuyến tính $(x_1, x_2, x_3, x_4, x_5) = (0, 1, 0, 6, 0)$ với mục tiêu tối ưu là $z = 6$. Nhận xét rằng ở từ vựng "khác" vừa xét các tỉ số $\frac{\bar{a}_{ik}}{\bar{b}_i}$ là $\frac{2}{2}$, $-\frac{0}{6}$ và $\frac{0}{0}$, không có tỉ số nào là $+\infty$. Chính giá trị $+\infty$ này sẽ gây rắc rối như ta thấy bây giờ.

Một phép xoay được gọi là *phép xoay suy biến* (degenerate pivot) nếu có tỉ số $\frac{\bar{a}_{ik}}{\bar{b}_i}$, khi xác định biến ra, là $+\infty$, tức là tử số dương và mẫu bằng 0. Khi đó ta cũng nói là từ vựng suy biến đã sinh ra phép xoay suy biến. Như vậy theo suy luận trên đây, từ vựng suy biến sẽ không gây khó khăn gì cho thuật toán đơn hình nếu nó không sinh ra phép xoay suy biến, còn nếu nó sinh ra phép xoay suy biến thì có thể phiền toái. Ta xét hai thí dụ sau đây.

THÍ DỤ 3.1.1 Xét từ vựng

$$\begin{array}{rcl}
z & = & -3 + 0{,}5x_1 - 2x_2 + 1{,}5w_1 \\
\hline
x_3 & = & 1 - 0{,}5x_1 \qquad\qquad - 0{,}5w_1 \\
w_2 & = & \qquad\quad x_1 - x_2 + \quad w_1.
\end{array}$$

Từ vựng này suy biến. Biến vào là x_2 và các tỉ số để xác định biến ra là 0 và $+\infty$. Vì cực đại trong đó là $+\infty$ nên biến ra là w_2. Nhưng, dễ thấy tỉ số tương ứng là $+\infty$ dẫn đến hiện tượng là x_2 vừa từ 0 trở thành dương thì w_2 trở thành âm, không chấp nhận được! Vì vậy mặc dù x_2 từ ngoài trở thành biến cơ sở nhưng từ vựng chấp nhận được tiếp theo có nghiệm tương ứng vẫn với $x_2 = 0$. Ta thấy điều đó sau khi thực hiện phép xoay (suy biến):

$$\begin{array}{rcl}
z & = & -3 - 1{,}5x_1 + 2w_2 - 0{,}5w_1 \\
\hline
x_3 & = & 1 - 0{,}5x_1 \qquad\qquad - 0{,}5w_1 \\
x_2 & = & \qquad\quad x_1 - w_2 + \quad w_1.
\end{array} \qquad (3.1)$$

Nhận xét rằng, chính do phép xoay là suy biến nên từ vựng mới này vẫn có nghiệm chấp nhận được $(x_1, x_2, x_3, w_1, w_2) = (0, 0, 1, 0, 0)$ và giá trị mục tiêu tương ứng $z = -3$ hoàn toàn như cũ. Nhưng dù sao từ vựng đã thay đổi, hi vọng là phép xoay tiếp theo sẽ cải thiện tình hình, nếu chưa thì phép xoay sau đó .... Cụ thể, ta thấy từ vựng (3.1) vẫn suy biến, nhưng x_1 là biến vào và x_3 là biến ra đã cho ta phép xoay không suy biến, dẫn đến từ vựng

$$\begin{array}{rcl}
z & = & -6 + 3x_3 + 2w_2 + w_1 \\
\hline
x_1 & = & 2 - 2x_3 \qquad\qquad - w_1 \\
x_2 & = & 2 - 2x_3 - w_2.
\end{array}$$

Từ vựng này là tối ưu.

Trong thực tế hơn nửa thế kỉ áp dụng phương pháp đơn hình người ta thấy thường gặp từ vựng suy biến nhưng sau một số bước lặp luôn có được phép xoay không suy biến làm thoát ra khỏi sự suy biến. Tuy nhiên về mặt lí thuyết, vẫn có thể xảy ra sự cố là sau một số phép xoay suy biến ta lại quay lại từ vựng suy biến đã gặp và thuật toán đơn hình lâm vào một vòng lặp vô hạn, không bao giờ tìm được nghiệm tối ưu. Trường hợp này gọi là *sự xoay vòng* (cycling).

Thí dụ đầu tiên về xoay vòng được tìm thấy ở [Hoffman 1953]. Ở nước ta, thí dụ xoay vòng đầu tiên là của GS Trần Vũ Thiệu (Viện Toán học, Hà Nội), năm 1967. [Marshall - Suurballe 1969] chứng minh rằng quy hoạch tuyến tính (chính tắc) xoay vòng phải có ít nhất sáu biến và ba ràng buộc.

THÍ DỤ 3.1.2 *Quy hoạch tuyến tính xoay vòng*. Xét quy hoạch

$$\begin{aligned} z &= -10x_1 + 57x_2 + 9x_3 + 24x_4 \\ \hline w_1 &= -0{,}5x_1 + 5{,}5x_2 + 2{,}5x_3 - 9x_4 \\ w_2 &= -0{,}5x_1 + 1{,}5x_2 + 0{,}5x_3 - x_4 \\ w_3 &= 1 - x_1. \end{aligned}$$

Như ở Mục 2.2 trình bày thuật toán đơn hình đã nói, có nhiều quy tắc xoay ứng với việc chọn biến vào và biến ra khác nhau, khi có nhiều "ứng viên" thỏa nguyên tắc chung của thuật toán đơn hình. Bây giờ ta chọn quy tắc xoay phổ biến nhất, tức là:

- chọn biến vào là biến có hệ số âm nhỏ nhất trong hàng mục tiêu của từ vựng;
- nếu có nhiều biến ra cùng ứng với tỉ số $\frac{\bar{a}_{ik}}{\bar{b}_i}$ lớn nhất thì chọn biến có chỉ số nhỏ nhất.

Sau sáu bước lặp như sau đây, thuật toán đơn hình quay trở lại từ vựng ban đầu:

$$\begin{array}{rcl}
z & = & 20w_1 - 53x_2 - 41x_3 + 204x_4 \\
\hline
x_1 & = & -2w_1 + 11x_2 + 5x_3 - 18x_4 \\
w_2 & = & w_1 - 4x_2 - 2x_3 + 8x_4 \\
w_3 & = & 1 + 2w_1 - 1x_2 - 5x_3 + 18x_4.
\end{array}$$

$$\begin{array}{rcl}
z & = & 6{,}75w_1 + 13{,}25w_2 - 14{,}5x_3 + 98x_4 \\
\hline
x_1 & = & 0{,}75w_1 - 2{,}75w_2 - 0{,}5x_3 + 4x_4 \\
x_2 & = & 0{,}25w_1 - 0{,}25w_2 - 0{,}5x_3 + 2x_4 \\
w_3 & = & 1 - 0{,}75w_1 - 13{,}25w_2 + 0{,}5x_3 - 4x_4.
\end{array}$$

$$\begin{array}{rcl}
z & = & -15w_1 + 93w_2 + 29x_1 - 18x_4 \\
\hline
x_3 & = & 1{,}5w_1 - 5{,}5w_2 - 2x_1 + 8x_4 \\
x_2 & = & -0{,}5w_1 + 2{,}5w_2 + x_1 - 2x_4 \\
w_3 & = & 1 - x_1.
\end{array}$$

$$\begin{array}{rcl}
z & = & -10{,}5w_1 + 70{,}5w_2 + 20x_1 + 9x_2 \\
\hline
x_3 & = & -0{,}5w_1 + 4{,}5w_2 + 2x_1 - 4x_2 \\
x_4 & = & -0{,}25w_1 + 1{,}25w_2 + 0{,}5x_1 - 0{,}5x_2 \\
w_3 & = & 1 - x_1.
\end{array}$$

$$\begin{array}{rcl}
z & = & 21x_3 - 24w_2 - 22x_1 + 93x_2 \\
\hline
w_1 & = & -2x_3 + 9w_2 + 4x_1 - 8x_2 \\
x_4 & = & 0{,}5x_3 - w_2 - 0{,}5x_1 + 1{,}5x_2 \\
w_3 & = & 1 - x_1.
\end{array}$$

$$
\begin{array}{rcl}
z & = & -10x_1 + 57x_2 + 9x_3 + 24x_4 \\
\hline
w_1 & = & -0{,}5x_1 + 5{,}5x_2 + 2{,}5x_3 - 9x_4 \\
w_2 & = & -0{,}5x_1 + 1{,}5x_2 + 0{,}5x_3 - x_4 \\
w_3 & = & 1 - x_1.
\end{array}
$$

Từ đây nếu ta vẫn áp dụng quy tắc "xoay" như trên, thuật toán đơn hình sẽ lặp lại vô hạn các vòng lặp gồm sáu bước như trên và không bao giờ dẫn đến nghiệm tối ưu.

Định lí sau đây khẳng định rằng thuật toán đơn hình luôn kết thúc được sau hữu hạn bước, chỉ trừ trường hợp bị xoay vòng.

ĐỊNH LÍ 3.1.1. ***Nếu thuật toán đơn hình kéo dài vô hạn thì nó phải xoay vòng.***

CHỨNG MINH. Một từ vựng được hoàn toàn xác định khi ấn định các biến nào là biến cơ sở, biến nào là không cơ sở. Có tất cả $C_{n+m}^{m} = \binom{n+m}{m}$ khả năng ấn định khác nhau. Nếu thuật toán kéo dài vô hạn nó phải gặp lại một từ vựng hơn một lần. Ngay từ khi nó lặp lại một từ vựng nào đó, ta phải có sự xoay vòng từ đây. □

CHÚ Ý. Tất cả các phép xoay trong một vòng, khi có hiện tượng xoay vòng, đều phải là phép xoay suy biến và các từ vựng trong vòng có cùng giá trị hàm mục tiêu và cùng nghiệm chấp nhận được tương ứng.

Quy hoạch tuyến tính xoay vòng hiếm đến nỗi hầu hết các thể nghiệm cụ thể phương pháp đơn hình trong thực tế, để giải các bài toán cỡ lớn là chính (các thuật toán, các phần mềm) đều không chú ý đến phòng ngừa xoay vòng. Tuy vậy, chỉ ra các nguyên tắc xoay để phương pháp đơn hình tránh được xoay vòng vẫn là quan trọng.

3.2 QUY TẮC XOAY BLAND

Trong [Bland 1977], quy tắc xoay, mà sau này gọi là quy tắc xoay Bland, rất đơn giản đã được chứng minh là luôn tránh được xoay vòng như sau.

Quy tắc Bland: Biến vào được chọn là biến có chỉ số nhỏ nhất trong các biến có hệ số mục tiêu âm; biến ra được chọn là biến có chỉ số nhỏ nhất trong số các biến có chỉ số $l \in B$ thỏa quy tắc chọn biến ra là

$$\frac{\bar{a}_{lk}}{\bar{b}_l} = \left(\max_{i \in B} \frac{\bar{a}_{ik}}{\bar{b}_i}\right).$$

ĐỊNH LÍ 3.2.1. *Thuật toán đơn hình sẽ kết thúc sau hữu hạn bước, tức là không bị xoay vòng, nếu áp dụng quy tắc xoay Bland.*

CHỨNG MINH. Giả sử phản chứng là có một vòng xoay gồm các từ vựng D_0, $D_1, \ldots, D_{k-1}, D_k$ với $D_k = D_0$. Ta gọi một biến là bất ổn (fickle) nếu nó là biến cơ sở trong một số từ vựng và lại là không cơ sở trong các từ vựng khác ở trong vòng xoay nói trên. Giả sử x_t là biến có chỉ số lớn nhất trong các biến bất ổn. Trong các từ vựng của vòng xoay phải có một từ vựng D mà tại đó x_t là biến ra và một biến bất ổn x_s khác là biến vào. Trong hai vòng xoay $D_0, D_1, \ldots, D_0, D_1, \ldots, D_0$ lại phải có từ vựng D_* mà ở đó x_t lại là biến vào. Giả sử D có dạng

$$z = v + \sum_{j \notin B} c_j x_j,$$

$$x_i = b_i - \sum_{j \notin B} a_{ij} x_j, \quad i \in B.$$

Đã biết là tất cả các phép xoay trong vòng xoay đều là suy biến và các từ vựng có cùng giá trị mục tiêu. Vì vậy phương trình mục tiêu của D_* có dạng

$$z = v + \sum_{j=1}^{n+m} c_j^* x_j. \qquad (3.2)$$

Ở đây ta viết tất cả các biến nhưng thực ra $c_j^* = 0$ nếu x_j là biến cơ sở trong D_*.

Quay lại xét từ vựng D ta thấy x_s là biến vào nên giá trị của nó tăng từ 0 lên $y > 0$. Theo phương pháp đơn hình trình bày ở Mục 2.2 ta thấy

$$x_s = y,\ x_j = 0 \text{ với } j \in N\backslash\{s\},\ x_i = b_i - a_{is}y \text{ với } i \in B. \qquad (3.3)$$

(Nhớ lại là B ký hiệu tập chỉ số các biến cơ sở và N là tập các chỉ số của các biến không cơ sở của từ vựng D.) Lúc này giá trị mục tiêu là $z = v + c_s y$. Vì các phương trình của D và D_* là tương đương, tức là có cùng tập nghiệm. (3.3) cũng là nghiệm cho D_* . Thay nó vào mục tiêu (3.2) ta được (vì giá trị mục tiêu của mọi từ vựng trong vòng lặp đều bằng nhau)

$$z = v + c_s^* y + \sum_{i \in B} c_i^*(b_i - a_{is}y) = v + c_s y.$$

Chuyển vế ta được

$$(c_s - c_s^* + \sum_{i \in B} c_i^* a_{is})y = \sum_{i \in B} c_i^* b_i.$$

Vì đẳng thức này đúng với mọi $y \le \left(\max\limits_{i \in B} \dfrac{\bar{a}_{ik}}{\bar{b}_i}\right)$ (xem lại việc chọn biến ra!), nên hệ số của y (và vế phải nữa) phải bằng 0:

$$c_s - c_s^* + \sum_{i \in B} c_i^* a_{is} = 0. \qquad (3.4)$$

Do x_s là biến vào ở từ vựng D nên $c_s < 0$. Vì x_t là biến bất ổn có chỉ số lớn nhất và x_s là biến bất ổn, nên $s < t$. Ở từ vựng D_* thì x_t là biến vào nên c_t^* là hệ số âm có chỉ số nhỏ nhất, tức là $c_t^* \ge 0$. Vì vậy, (3.4) suy ra $\sum\limits_{i \in B} c_i^* a_{is} > 0$, tức là phải có chỉ số $r \in B$ để $c_r^* a_{rs} > 0$. Từ đó, $c_r^* \neq 0$ và x_r là biến không cơ sở của từ vựng D_*. Mà $r \in B$ nghĩa là x_r là biến cơ sở của D. Vậy x_r là biến bất ổn và $r \leq t$. Nhận xét rằng $c_t^* a_{ts} < 0$ vì $c_t^* < 0$ do x_t là biến vào trong D_* còn $a_{ts} > 0$ vì x_t là biến ra trong D. Mà $c_r^* a_{rs} < 0$ nên $r \neq t$. Do đó $r < t$. Do vậy $c_r^* \geq 0$, vì x_t đã là biến vào của D_*. Từ $c_r^* a_{rs} > 0$ ta thấy $a_{rs} > 0$.

Chú ý rằng mọi từ vựng trong vòng xoay tương ứng cùng một nghiệm chấp nhận được, ta thấy mọi biến bất ổn đều phải bằng 0. Nói riêng $x_r = 0$. Nhưng trong D thì x_r là biến cơ sở, nên $b_r = 0$. Do $a_{rs} > 0$ và $b_r = 0$ nên x_r phải là ứng viên của biến ra trong từ vựng D; mà $r < t$ nên theo quy tắc Bland nó phải được chọn trước x_t. Ta đã gặp mâu thuẫn. □

3.3 PHƯƠNG PHÁP NHIỄU LOẠN

Phương pháp chống xoay vòng đầu tiên là *phương pháp nhiễu loạn* (perturbation method) được đề xướng bởi A. Orden và được phát biểu (độc lập với Orden) ở dạng hoàn thiện bởi [Charnes 1952]. Đây cũng là cơ sở trực quan để hình thành *phương pháp từ điển* (lexicographic method hoặc lexical method) trong [Dantzig - Orden - Wolfe 1955].

Nhận xét quyết định dẫn đến ý tưởng của phương pháp nhiễu loạn là suy biến là hiện tượng rất "tình cờ". Các thí dụ của chúng ta trong sách đều chỉ dùng đến các hệ số nguyên và nhỏ, nên ta không ngạc nhiên lắm khi thực hiện bước lặp đơn hình, tức là chỉ thực hiện các phép toán đại số, ta gặp hiện tượng "khử lẫn nhau" làm cho có $\bar{b}_i$ trở thành số 0 và dẫn đến từ vựng suy biến. Thực ra các hệ số là các số thực, tức là có rất nhiều hoặc thậm chí vô hạn chữ số thập phân. Hiện tượng "khử lẫn nhau" để xuất hiện số 0 mới là gần như không xảy ra. Nói một cách khác, nếu ta đang gặp một từ vựng suy biến, thì chỉ làm nhiễu các vế phải $\bar{b}_k$, $k = 1, ..., m$, đi rất nhỏ ta sẽ được một từ vựng không suy biến. Khi nhiễu là đủ nhỏ thì nghiệm tối ưu của bài toán nhiễu thực tế có thể dùng như nghiệm tối ưu của bài toán gốc.

Trong một số phần mềm máy tính của thuật toán đơn hình, người ta thay các b_k, $k = 1, ..., m$, bởi $b_k + \varepsilon$, với ε nhỏ ($\varepsilon = 10^{-6}$ chẳng hạn) và gần như không thấy gặp phép xoay suy biến nữa. Tuy nhiên, đây chưa phải là cách chống xoay vòng tuyệt đối. Chẳng hạn, xét thí dụ nhiễu lên quy hoạch tuyến tính xoay vòng tương tự Thí dụ 3.1.2:

$$\begin{aligned}
\min\ & -10x_1 + 57x_2 + 9x_3 + 24x_4 + 100x_5,\\
& 0{,}5x_1 - 5{,}5x_2 - 2{,}5x_3 + 9x_4 + x_5 \le 1 + \varepsilon,\\
& 0{,}5x_1 - 1{,}5x_2 - 0{,}5x_3 + x_4 + x_5 \le 1 + \varepsilon,\\
& x_1 + x_5 \le 1 + \varepsilon,\\
& x_5 \le 1 + \varepsilon,\\
& x_1, x_2, \ldots, x_5 \ge 0.
\end{aligned}$$

Sau ngay bước lặp từ vựng đầu tiên ta gặp từ vựng suy biến

$$\begin{array}{l}
z = -100 - 100\varepsilon - 10x_1 + 57x_2 + 9x_3 + 24x_4 - 100x_9\\
\hline
x_6 = \quad - 0{,}5x_1 + 5{,}5x_2 + 2{,}5x_3 - 9x_4 + x_9\\
x_7 = \quad - 0{,}5x_1 + 1{,}5x_2 + 0{,}5x_3 - x_4 + x_9\\
x_8 = 1 - x_1 + x_9\\
x_5 = 1 + \varepsilon - x_9.
\end{array}$$

Độc giả có thể làm tiếp thuật toán đơn hình và nhận được xoay vòng sau 6 bước lặp nữa. Nguyên nhân của sự suy biến và xoay vòng này là vì các số ε đã "khử" lẫn nhau ngay ở bước lặp đầu tiên. Để tránh sự "khử" này ta cần nhiễu các vế phải $b_1, ..., b_m$ các lượng $\varepsilon_1, ..., \varepsilon_m$ nhỏ nhưng "khác hẳn nhau". Nói chính xác hơn, ta chọn

$$0 << \varepsilon_m << \varepsilon_{m-1} << \dots << \varepsilon_1 << 1,$$

(dấu << có nghĩa là cấp nhỏ hơn hẳn). Chẳng hạn người ta hay lấy $\varepsilon_1 = \varepsilon, \dots, \varepsilon_m = \varepsilon^m$ (các lũy thừa).

Để chứng minh Định lí 3.3.1 dưới đây về phương pháp nhiễu loạn nói riêng cũng như để nghiên cứu định tính và thuật toán cho quy hoạch tuyến tính ta đưa ra một số khái niệm và cách biểu diễn khác với từ vựng. Trước hết, cần nhớ rằng khi trình bày thuật toán đơn hình, ta luôn xét quy hoạch tuyến tính dạng chính tắc

$$\begin{aligned} \min z &= c^T x, \\ Ax &= b, \\ x &\geq 0, \end{aligned} \tag{3.5}$$

ở đây A là ma trận cấp m x n, với m < n. Không mất tổng quát ta luôn giả thiết rank A = m, vì nếu có một hàng là tổ hợp tuyến tính của các hàng khác thì ta chỉ việc bỏ ràng buộc ứng với hàng đó đi để nhận được ràng buộc tương đương có các hàng độc lập tuyến tính.

Nghiệm cơ sở của hệ $Ax = b$ là nghiệm $(x_1, \dots, x_n)$ nhận được khi cho n − m thành phần x_j bằng 0 và giải ra được m thành phần còn lại. Khi đó n − m thành phần bằng 0 là n − m biến không cơ sở x_j, $j \in N$, và m biến còn lại x_i, $i \in B$, là các biến cơ sở. Như vậy từ vựng ứng với một nghiệm cơ sở xác định khi $Ax = b$ đã viết ở dạng đặc biệt là

$$x_i = \bar{b}_i - \sum_{j \in N} \bar{a}_{ij} x_j \quad , i \in B. \tag{3.6}$$

Tổng quát, một cách tự nhiên ta gọi m cột A_i, $i \in B$, của ma trận A là cơ sở của nghiệm cơ sở đã xác định đó. Nhớ rằng các cột này là độc lập tuyến tính vì với dạng từ vựng (3.6) thì ma trận A đã có dạng đặc biệt mà các cột này là các vectơ đơn vị $(0, \dots, 0, 1, 0, \dots, 0)^T$. Nếu nghiệm cơ sở ứng với từ vựng là không suy biến, tức là từ vựng không suy biến ($\bar{b}_i \neq 0 \ \forall i \in B$), thì tập chỉ số

$$I := \{\, i : x_i \neq 0 \,\} \tag{3.7}$$

bằng B. Trái lại nếu nghiệm cơ sở là suy biến thì $I \subset B$ và $I \neq B$. Nếu nghiệm cơ sở là không suy biến thì chỉ có đúng n − m thành phần bằng 0, nên chỉ có thể ứng với một từ vựng duy nhất. Nếu nghiệm cơ sở là suy biến, tức là có hơn n − m thành phần 0, thì có nhiều cách chọn n − m thành phần 0 làm biến không cơ sở, miễn là m thành phần còn lại x_i ứng với các cột A_i độc lập tuyến tính của ma trận A, để có thể giải được ra các x_i này theo n − m thành phần x_j đã chọn làm biến không cơ sở. Do đó ta cần định nghĩa tổng quát khái niệm cơ sở như sau.

ĐỊNH NGHĨA 3.3.1. *Cơ sở* (basis) của nghiệm cơ sở $(x_1^0, ..., x_n^0)$ là một hệ m vectơ cột $\{A_i: i \in B\}$ độc lập tuyến tính của ma trận A sao cho

$$B \supset I := \{i: x_i^0 \neq 0\}.$$

Như vậy cơ sở được xác định không cần thông qua từ vựng. Ta cũng phát biểu lại cụ thể nghiệm cơ sở suy biến như sau, không phụ thuộc vào từ vựng. Nghiệm cơ sở được gọi là *nghiệm cơ sở không suy biến* nếu có đúng n − m thành phần bằng không. Nghiệm cơ sở là *suy biến* nếu có nhiều hơn n − m thành phần bằng 0. Nếu nghiệm cơ sở là không suy biến thì nó có một cơ sở duy nhất là $\{A_i: i \in I\}$ và ứng với một từ vựng duy nhất. Nếu nghiệm cơ sở là suy biến thì nó có thể có nhiều cơ sở và ứng với nhiều từ vựng khác nhau. Chẳng hạn, ta đã biết là từ vựng nhận được bằng phép xoay suy biến từ một từ vựng suy biến thì vẫn xác định nghiệm cơ sở đó. Nói cách khác, mọi từ vựng trong vòng xoay cũng ứng với một nghiệm cơ sở.

Rõ ràng thay cho $b = Ax^0$ ta có thể viết $b = \sum_{i \in I} x_i^0 A_i$.

Ta gọi một nghiệm cơ sở là *nghiệm cơ sở chấp nhận được* nếu nó thỏa cả mọi ràng buộc dấu $x_j \geq 0$, $j = 1, ..., n$.

Ta nói $P := \{x \in \mathbb{R}^n: Ax = b, x \geq 0\}$ là *tập lồi đa diện không suy biến* (nondegenerate polyhedron) nếu mọi nghiệm cơ sở của nó đều không suy biến.

ĐỊNH LÍ 3.3.1. Xét nhiễu loạn nhận được bằng cách nhiễu vế phải b thành

$$b(\varepsilon) := b + \sum_{k=1}^{m} \varepsilon^k C_k ,$$

ở đây C_k, $k = 1, ..., m$, là m vectơ độc lập tuyến tính tùy ý. Khi đó luôn tồn tại $\varepsilon_0 > 0$ để miền ràng buộc của bài toán nhiễu $P(\varepsilon) := \{x \in \mathbf{R}^n: Ax = b(\varepsilon), x \geq 0\}$ là tập lồi đa diện không suy biến với mọi $\varepsilon \in (0, \varepsilon_0)$.

CHỨNG MINH. Với $\{A_i: i \in B\}$ là m cột độc lập tuyến tính bất kì của A, gọi $x(\varepsilon)$ và x là nghiệm cơ sở tương ứng của bài toán nhiễu và không nhiễu, tức là

$$\sum_{i \in B} x_i(\varepsilon) A_i = b + \sum_{k=1}^{m} \varepsilon^k C_k,$$
$$x_j(\varepsilon) = 0, j \in \{1, ..., n\} \setminus B; \tag{3.8}$$

và

$$\sum_{i \in B} x_i A_i = b,$$
$$x_j = 0, j \in \{1, ..., n\} \setminus B. \tag{3.9}$$

Khi đó $x(\varepsilon)$ và x đều xác định duy nhất. Các vectơ C_k cũng biểu diễn được duy nhất qua $\{A_i: i \in B\}$:

$$C_k = \sum_{i \in B} \gamma_{ik} A_i \ , k = 1, ..., m. \tag{3.10}$$

Thay (3.9) và (3.10) vào (3.8) ta được

$$\sum_{i \in B} x_i(\varepsilon) A_i = \sum_{i \in B} x_i A_i + \sum_{k=1}^{m} \varepsilon^k \sum_{i \in B} \gamma_{ik} A_i = \sum_{i \in B} (x_i + \sum_{k=1}^{m} \gamma_{ik} \varepsilon^k) A_i.$$

Do các A_i độc lập tuyến tính ta có

$$x_i(\varepsilon) = x_i + \sum_{k=1}^{m} \gamma_{ik}\varepsilon^k \quad, i \in B.$$

Với mỗi $i \in B$, $x_i(\varepsilon)$ là một đa thức cấp $\leq m$ của ε. Gọi e_i là nghiệm dương nhỏ nhất của phương trình $x_i(\varepsilon) = 0$ (nếu không có nghiệm dương thì đặt $e_i = +\infty$). Ta sẽ chứng minh $\varepsilon_0 := \min_{i \in B} e_i$ là số cần tìm. Giả sử $\varepsilon \in (0, \varepsilon_0)$ và $x(\varepsilon)$ là một nghiệm cơ sở của $P(\varepsilon)$. Khi đó $x(\varepsilon)$ phải ứng một cơ sở $\{A_i: i \in B\}$ nào đó. Vậy $x(\varepsilon)$ thỏa hệ dạng (3.8) và $x_i(\varepsilon) \geq 0$ $\forall i \in B$. Giả sử phản chứng là $x(\varepsilon)$ suy biến, tức là $\exists i \in B$ để $x_i(\varepsilon) = 0$. Khi đó ε là nghiệm phương trình $x_i(\varepsilon) = 0$, mâu thuẫn với định nghĩa ε_0. Vậy mọi nghiệm cơ sở của $P(\varepsilon)$ là không suy biến.□

Theo Định lí 3.3.1 ta có thể dùng quy tắc xoay bất kì (cụ thể là ta được tự do khi có nhiều ứng viên cho biến vào cũng như biến ra) để thực hiện thuật toán đơn hình cho bài toán nhiễu mà không bao giờ gặp từ vựng suy biến (vì không tồn tại từ vựng suy biến của bài toán nhiễu). Nếu ε là nhỏ, có thể coi nghiệm tối ưu của bài toán nhiễu là nghiệm tối ưu của bài toán ban đầu. Nói chính xác hơn là cho $\varepsilon \to 0^+$ ta được nghiệm tối ưu cần tìm.

3.4 PHƯƠNG PHÁP TỪ ĐIỂN

Phương pháp từ điển của [Dantzig - Orden - Wolfe 1955] được xây dựng dựa vào phương pháp nhiễu loạn. Ý tưởng của phương pháp này là chọn ứng viên tùy ý cho biến vào (miễn là có hệ số mục tiêu âm), còn biến ra thì chọn dựa theo phương pháp nhiễu loạn, nhưng không cần nhiễu mà nhận xét ngay trên các hệ số của bài toán cần xét để chọn. Tất nhiên việc nhận xét này là dựa theo phương pháp nhiễu loạn.

Ta cần định nghĩa sau. Ta nói vectơ $a \neq 0$ là *dương theo từ điển* (lexically positive), kí hiệu là $a \succ 0$ nếu thành phần khác 0 đầu tiên của a là dương. Ta nói vectơ a lớn hơn vectơ b theo từ điển, và viết $a \succ b$ nếu $a - b \succ 0$.

BỔ ĐỀ 3.4.1. *Cho các đa thức của* ε

$$f_j(\varepsilon) = \sum_{k=0}^{m} p_{jk}\varepsilon^k \quad , j \in J.$$

Gọi p^j là vectơ hệ số của đa thức f_j, tức là $p^j = (p_{j1}, ..., p_{jm})$ và p^l là cực tiểu theo từ điển của p^j, $j \in J$. Khi đó tồn tại ε_0 để $\forall \varepsilon \in (0, \varepsilon_0)$, $\forall j \neq l$, ta có $f_j(\varepsilon) > f_l(\varepsilon)$.

CHỨNG MINH. Ta có

$$f_j(\varepsilon) - f_l(\varepsilon) = \sum_{k=0}^{m}(p_{jk} - p_{lk})\varepsilon^k . \tag{3.11}$$

Mà $p^j - p^l \succ 0$ nên hệ số khác 0 đầu tiên ở vế phải (3.11) là dương. Do đó sẽ có $\delta_j > 0$ đủ bé để với mọi $\varepsilon \in (0, \delta_j)$ thì $f_j(\varepsilon) - f_l(\varepsilon) > 0$. Lấy $\varepsilon_0 = \min\limits_{j \neq l}\delta_j$ ta được số cần tìm. □

Bổ đề trên nói rằng việc so sánh các đa thức của ε với các giá trị đủ bé của ε là tương đương với so sánh các vectơ hệ số của đa thức theo *thứ tự từ điển* (lexical ordering).

Để xây dựng phương pháp từ điển ta viết quy hoạch tuyến tính ở dạng (2.9), hoặc chỉ viết các hệ số như bảng (2.10). (Nhận xét rằng (2.9) chẳng qua là từ vựng nhưng đã chuyển vế biến cơ sở sang cùng bên với biến không cơ sở và tách ra thành m cột.) Như vậy ta có ngay cơ sở $\{A_i: i \in B\}$ ứng với nghiệm cơ sở (của từ vựng này) gồm m vectơ đơn vị $A_i = (0,...,0,1,0,...,0)^T$. Do đó, để thuận tiện ta đặt lại kí hiệu các biến, coi các biến bù w_i là các biến đầu tiên $x_1, ..., x_m$ và khi đó $B = \{1, ..., m\}$. (Khi thực hiện dạng bảng có thể ta vẫn để các cột A_i cơ sở này ở cuối bảng như (2.10).)

Song song với bài toán chính ta xét bài toán nhiễu loạn với $b(\varepsilon) = b + \sum_{i=1}^{m} \varepsilon^i A_i$, tức là ta đã lấy các vectơ C_i ở Định lí 3.3.1 là $C_i = A_i$.

Chẳng hạn bài toán nhiễu loạn của quy hoạch (2.9) là

$$
\begin{array}{rcl}
-z - 5x_4 - 4x_5 - 3x_6 & = & 0 \\
2x_4 + 3x_5 + x_6 + x_1 & = & 5 + \varepsilon \\
4x_4 + x_5 + 2x_6 + x_2 & = & 11 + \varepsilon^2 \\
3x_4 + 4x_5 + 2x_6 + x_3 & = & 8 + \varepsilon^3.
\end{array}
\tag{3.12}
$$

Chú ý là ta đổi chỉ số các biến. Đồng thời, nhận xét quan trọng nhất ở đây là các vectơ cột ở phần nhiễu hoàn toàn trùng với m vectơ cột cuối bên vế trái. Do đó khi thực hiện các phép toán hàng (tức là khi thực hiện thuật toán đơn hình) ta vẫn luôn có sự trùng nhau này. Gọi $x(\varepsilon)$ và x là nghiệm cơ sở tương ứng của bài toán nhiễu và bài toán ban đầu. Khi đó các phương trình ràng buộc sẽ là

$$\sum_{k=1}^{m} x_k(\varepsilon) A_k = b + \sum_{i=1}^{m} \varepsilon^i A_i, \tag{3.13}$$

$$x_j(\varepsilon) = 0,\ j = m+1, \dots, n.$$

Đây là biểu diễn ở bước đầu tiên, khi cơ sở là $\{A_k : k = 1, \dots, m\}$ theo cách chọn chỉ số của ta. Ở các bước lặp sau, cơ sở $\{A_k : k \in B\}$ sẽ ứng với tập chỉ số B khác $\{1, \dots, m\}$. Nhưng, ta có quan hệ tương tự (3.13) chỉ thay cho $\sum_{k=1}^{m}$ ở vế trái là $\sum_{i \in B}$ và thay cho $j = m+1, \dots, n$ là $j \in N$. Do đó, để xác định, sau đây ta lí luận theo chỉ số của bước đầu.

Khi chọn biến ra cho quy hoạch nhiễu loạn, sau khi đã có biến vào là x_j, $j \in \{m+1, \dots, n\}$, ta phải lấy chỉ số l sao cho

$$\frac{\bar{a}_{lj}}{\bar{b}_l(\varepsilon)} = \max_{k \in B=\{1,\dots,m\}} \frac{\bar{a}_{kj}}{\bar{b}_k(\varepsilon)} = \max_k \frac{\bar{a}_{kj}}{\bar{b}_k + \sum_{i=1}^{m} \bar{a}_{ki}\,\varepsilon^i}.$$

Tương đương, l là chỉ số sao cho

$$\frac{\bar{b}_l(\varepsilon)}{\bar{a}_{lj}} = \min_{k \in \{1,\dots,m\}} \frac{\bar{b}_k}{\bar{a}_{kj}} + \sum_{i=1}^{m} \frac{\bar{a}_{ki}}{\bar{a}_{kj}}\,\varepsilon^i. \tag{3.14}$$

Theo Bổ đề 3.4.1, với ε đủ nhỏ việc so sánh m đa thức của ε ở vế phải (3.14) là tương đương với việc so sánh theo từ điển các vectơ hệ số của chúng

$$p^k = \left(\frac{\bar{b}_k}{\bar{a}_{kj}}, \frac{\bar{a}_{k1}}{\bar{a}_{kj}}, \dots, \frac{\bar{a}_{km}}{\bar{a}_{kj}}\right), \ k = 1, \dots, m. \tag{3.15}$$

Chú ý rằng $\bar{a}_{kj}$ là phần tử trên cột xoay j (ở quy hoạch (3.12) là j = 4 nếu ta chọn biến vào là biến có hệ số mục tiêu âm, nhỏ nhất). Còn $(\bar{a}_{k1}, \dots, \bar{a}_{km})$ là các hệ số của $\varepsilon^1, \dots, \varepsilon^m$ ở hàng k vế phải, nhưng cũng chính là m số cuối của hàng k (ứng với m cột cơ sở ban đầu). Chẳng hạn ở quy hoạch tuyến tính (3.12) ta phải tìm cực tiểu theo từ điển của các vectơ

$$\begin{gathered}\left(\frac{5}{2}, \frac{1}{2}, \frac{0}{2}, \frac{0}{2}\right),\\ \left(\frac{11}{4}, \frac{0}{4}, \frac{1}{4}, \frac{0}{4}\right),\\ \left(\frac{8}{3}, \frac{0}{3}, \frac{0}{3}, \frac{1}{3}\right)\end{gathered} \tag{3.16}$$

và ta được vectơ đầu, tức là $l = 1$: x_1 là biến ra.

Như vậy việc chọn biến ra ở quy hoạch nhiễu loạn tương đương với việc tìm cực tiểu theo từ điển các vectơ (3.15) của bài toán không nhiễu. Tức là ta không cần viết bài toán nhiễu loạn nữa.

Chú ý rằng, khi áp dụng thuật toán đơn hình lên quy hoạch tuyến tính dạng (3.12) có nhiễu loạn và không nhiễu, thì ở các bước ta nhận được kết quả hoàn toàn như nhau cho hai bài toán, điều khác duy nhất chỉ là ở "phần nhiễu ε" ở vế phải. Nhưng hệ số m của cột này hoàn toàn trùng với hệ số của m cột cuối ở vế trái. Vì thế ta không cần viết bài toán nhiễu loạn nữa.

Ở Mục 3.3 ta đã biết phương pháp nhiễu loạn không thể gặp xoay vòng, nên phương pháp từ điển cũng không thể gặp xoay vòng.

NHẬN XÉT.

1. Luôn có vectơ cực tiểu theo từ điển duy nhất ở (3. 15). Thật vậy, nếu có hai vectơ tỉ lệ $p^{k_1} = \gamma p^{k_2}$, thì có nghĩa là hai hàng trong ma trận $(A_1, \ldots, A_m)$ là tỉ lệ, mâu thuẫn với việc các vectơ A_k, $k = 1, \ldots, m$, là cơ sở nên độc lập tuyến tính.

2. Theo thuật toán đơn hình trình bày ở Mục 2.2 ta đã biết, khi tìm $\max\limits_{k \in B} \dfrac{\bar{a}_{kj}}{\bar{b}_k}$ để chọn biến ra ta chỉ xét các $k \in B$ mà $\bar{a}_{kj} \geq 0$ (nếu mọi $a_{kj} < 0$ thì, theo Mục 2.4, quy hoạch tuyến tính không giới nội, tức là không có nghiệm tối ưu). Như vậy, có thể coi là ở phương pháp từ điển ta so sánh theo từ điển mọi vectơ (3. 15) có $\bar{a}_{kj} \geq 0$ để chọn biến ra. Song cũng có thể coi là ta chỉ so sánh các ứng viên thôi, tức là các vectơ ở (3. 15) với $\bar{a}_{kj} \geq 0$ và $\dfrac{\bar{b}_k}{\bar{a}_{kj}}$ (tức là thành phần đầu tiên) là lớn nhất thôi.

3. Trong áp dụng thực tế ta không cần dùng các quy tắc xoay để tránh xoay vòng ngay từ đầu thuật toán. Chẳng hạn, với bài toán cỡ lớn, chỉ sau khi bị khoảng năm chục phép xoay suy biến (bước lặp suy biến) ta mới áp dụng quy tắc Bland, phương pháp nhiễu hoặc từ điển.

BÀI TẬP CHƯƠNG 3

3.1. Thực hiện và so sánh phương pháp đơn hình với hai cách xoay khác nhau: quy tắc hệ số lớn nhất Dantzig và quy tắc chỉ số bé nhất (Bland) cho các quy hoạch sau

(a)
$$\begin{aligned} &\max 4x_1 + 5x_2,\\ &2x_1 + 2x_2 \le 9,\\ &x_1 \le 4,\\ &x_2 \le 3,\\ &x_1, x_2 \ge 0; \end{aligned}$$

(b)
$$\begin{aligned} &\min -2x_1 - x_2,\\ &3x_1 + x_2 \le 3,\\ &x_1, x_2 \ge 0; \end{aligned}$$

(c)
$$\begin{aligned} &\max 3x_1 + 5x_2,\\ &x_1 + 2x_2 \le 5,\\ &x_1 \le 3,\\ &x_2 \le 2,\\ &x_1, x_2 \ge 0. \end{aligned}$$

3.2. Chứng minh rằng nếu nghiệm chấp nhận được cơ sở x có hai cơ sở tương ứng thì x là nghiệm cơ sở suy biến. Mệnh đề ngược có đúng không?

3.3. Chứng tỏ rằng quy hoạch tuyến tính sau là xoay vòng

$$\begin{aligned} &\min -10x_1 + 57x_2 + 9x_3 + 24x_4 - 100x_5,\\ &0{,}5x_1 - 5{,}5x_2 - 2{,}5x_3 + 9x_4 + x_5 \le 1,\\ &0{,}5x_1 - 1{,}5x_2 - 0{,}5x_3 + x_4 + x_5 \le 1,\\ &x_1 + x_5 \le 1,\\ &x_5 \le 1,\\ &x_1, \dots, x_5 \ge 0. \end{aligned}$$

Chỉ ra rằng nếu nhiễu các vế phải từ 1 thành $1 + \varepsilon$ thì bài toán nhiễu vẫn bị xoay vòng. Cuối cùng, giải bài toán bằng cách nhiễu các vế phải là $1 + \varepsilon$, $1 + \varepsilon^2$, $1 + \varepsilon^3$, $1 + \varepsilon^4$.

3.4. Giải bằng phương pháp nhiễu loạn và bằng quy tắc Bland

$$\begin{aligned}
\min\ -10x_1 + 57x_2 + 9x_3 + 24x_4&,\\
0{,}5x_1 - 5{,}5x_2 - 2{,}5x_3 + 9x_4 &\le 0,\\
0{,}5x_1 - 1{,}5x_2 - 0{,}5x_3 + x_4 &\le 0,\\
x_1 &\le 1,\\
x_1, x_2, x_3, x_4 &\ge 0.
\end{aligned}$$

3.5. Giả sử một quy hoạch tuyến tính có tính chất sau: khi giải bằng phương pháp đơn hình bao giờ cũng chỉ gặp một ứng viên duy nhất cho biến ra.

(a) Bài toán này có thể có từ vựng suy biến không? tại sao?

(b) Bài toán có thể bị xoay vòng không? tại sao?

3.6. Giải bằng thuật toán đơn hình và bằng quy tắc Bland

$$\begin{aligned}
\min\ 2x_1 - 3x_2 + 4x_3&,\\
-2x_2 - 3x_3 &\ge -5,\\
x_1 + x_2 + 2x_3 &\le 4,\\
x_1 + 2x_2 + 3x_3 &\le 7,\\
x_1, x_2, x_3 &\ge 0.
\end{aligned}$$

3.7. Xét quy hoạch

$$\begin{aligned}
\max\ 2x_1 + 3x_2 - x_3 - 12x_4&,\\
-2x_1 - 9x_2 + x_3 + 9x_4 &\le 0,\\
\frac{1}{3}x_1 + x_2 - \frac{1}{3}x_3 - 2x_4 &\le 0,\\
x_1, x_2, x_3, x_4 &\ge 0.
\end{aligned}$$

Chứng tỏ rằng nếu ta thực hiện phương pháp đơn hình theo quy tắc xoay "nếu có nhiều ứng viên cho hàng xoay ta chọn hàng cao nhất", thì bài toán bị xoay vòng.

Hãy giải bằng phương pháp nhiễu loạn và phương pháp từ điển.

3.8. Xét quy hoạch

$$\begin{aligned}
\min\ & 4x_1 - 6x_5 - 5x_6 + 64x_7,\\
& x_1 + \frac{1}{3}x_4 - 2x_5 - x_6 + 12x_7 = 0,\\
& x_2 + \frac{1}{2}x_4 - x_5 - \frac{1}{6}x_6 + \frac{2}{3}x_7 = 0,\\
& x_3 + x_5 + x_6 - 9x_7 = 2,\\
& x_j \geq 0,\ j = 1,\ldots, 7.
\end{aligned}$$

(a) Giải bằng hai cách: quy tắc Bland và phương pháp từ điển.

(b) Chứng tỏ là bài toán bị xoay vòng khi áp dụng phương pháp đơn hình (với một quy tắc xoay nào đó, thử tìm!)

3.9. Xét từ vựng

$$\begin{aligned}
z &= 5 + 2x_2 - 2x_3 + 3x_5\\
\hline
x_6 &= 4 - 2x_2 - x_3 + x_5\\
x_4 &= 2 - x_2 + x_3 - x_5\\
x_1 &= 6 - 2x_2 - 2x_3 - 3x_5.
\end{aligned}$$

(a) Chỉ ra tất cả các cặp (x_r, x_s) mà x_r có thể là biến vào và x_s có thể là biến ra.

(b) Chỉ ra tất cả các cặp như trên nếu dùng quy tắc hệ số lớn nhất để chọn biến vào.

(c) Chỉ ra tất cả các cặp như trên nếu dùng quy tắc Bland để chọn biến vào và ra.

3.10. Chứng minh rằng nếu $n - m = 2$ (trong quy hoạch chính tắc) thì phương pháp đơn hình không thể bị xoay vòng, dù với quy tắc xoay nào.

3.11. Tìm một thí dụ với $n - m = 3$ và một quy tắc xoay sao cho phương pháp đơn hình là xoay vòng.

3.12. Xét quy hoạch tuyến tính $\min c^T x$, $x \in P$, ở đây $P \in R^n$ là một tập lồi đa diện giới nội. Đặt

$$Q = \{(tx, t) \in R^{n+1} : x \in P, t \in [0, 1]\}.$$

(a) chứng tỏ rằng Q là một tập lồi đa diện.

(b) Cho một thí dụ bằng hình học tập P và Q, với n = 2, sao cho điểm 0 $\in R^{n+1}$ là nghiệm chấp nhận được cơ sở suy biến của Q.

(c) Chứng minh rằng điểm $0 \in R^{n+1}$ làm cực tiểu $(c, 0)^T y$ trên tập $y \in Q$ nếu và chỉ nếu giá trị mục tiêu tối ưu của quy hoạch tuyến tính ban đầu là không âm.

Chú ý rằng bài tập này xét vấn đề "khi nào nghiệm cơ sở chấp nhận được suy biến là tối ưu" và chỉ ra rằng về bản chất trả lời câu hỏi này cũng khó như giải một quy hoạch tuyến tính tổng quát.

3.13. Giả sử $0 < \varepsilon_m << \varepsilon_{m-1} << \ldots << \varepsilon_1 << 1$.

(a) Xếp theo thứ tự từ điển các đại lượng sau: $5 - \varepsilon_1$, $3 + 10\varepsilon_1$, 3, $3 - 2\varepsilon_1 + 3\varepsilon_2$, $\varepsilon_2 + 2\varepsilon_3$, $3 + 4\varepsilon_1 - \varepsilon_3$, $3 - 4\varepsilon_1 + \varepsilon_2 + \varepsilon_3$.

(b) Giả sử $r = \sum_{i=1}^{m} r_i\varepsilon^i$, $s = \sum_{i=1}^{m} s_i\varepsilon_i$ và $t = \sum_{i=1}^{m} t_i\varepsilon_i$ ở đây $\varepsilon_0 = 0$. Chứng minh tính chất bắc cầu: nếu $r \succ s$ và $s \succ t$, thì $r \succ t$.

(c) Chứng minh rằng trong một tập hữu hạn các biểu thức theo ε_i như ở (b) luôn có biểu thức là cực tiểu theo từ điển và biểu thức là cực đại theo từ điển.

(d) Chứng minh rằng với mọi cặp r và s như ở (b) sẽ có một số $\delta > 0$ sao cho hai khẳng định sau là tương đương

(i) $r \prec s$ ($\prec$ là nhỏ hơn theo từ điển);

(ii) với mọi bộ số $\varepsilon_1, \ldots, \varepsilon_m$ sao cho

$$0 < \varepsilon_1 < \delta \text{ và } 0 < \varepsilon_i < \delta\varepsilon_i - 1, \ i = 2, \ldots, m,$$

$r < s$ ($<$ là nhỏ hơn theo so sánh số thực).

3.14. Giải

(a)
$$\begin{aligned} \max\ & x_1 + 3x_2 - x_3, \\ & 2x_1 + 2x_2 - x_3 \le 10, \\ & 3x_1 - 2x_2 + x_3 \le 10, \\ & x_1 - 3x_2 + x_3 \le 10, \\ & x_1,\ x_2,\ x_3 \ge 0; \end{aligned}$$

(b)
$$\begin{aligned} \max\ & 3x_1 + x_2, \\ & -x_1 + x_2 \ge 1, \\ & -x_1 - x_2 \le -3, \\ & 2x_1 + x_2 \le 4, \\ & x_1,\ x_2 \ge 0; \end{aligned}$$

(c)
$$\begin{aligned} \min\ & -3x_1 - x_2, \\ & x_1 - x_2 \le -1, \\ & x_1 + x_2 \ge 3, \\ & 2x_1 + x_2 \le 2, \\ & x_1, x_2 \ge 0; \end{aligned}$$

(d)
$$\begin{aligned} \max\ & 3x_1 + x_2, \\ & x_1 - x_2 \le -1, \\ & -x_1 - x_2 \le -3, \\ & 2x_1 - x_2 \le 2, \\ & x_1, x_2 \ge 0; \end{aligned}$$

(e)
$$\begin{aligned} \min\ & -5x_1 + 3x_2 - 2x_3 + x_4 - 2x_5, \\ & x_1 + x_2 + 3x_3 + 2x_4 + x_5 \le 1000, \\ & 5x_1 + 3x_2 + x_3 + 5x_4 + 2x_5 \le 2000, \\ & x_1, \ldots, x_5 \ge 0; \end{aligned}$$

(f)
$$\begin{aligned} \min\ & 2x_1 - 5x_2, \\ & x_1 + 3x_2 \le 10, \\ & 2x_1 - 3x_2 \le 0, \\ & -x_1 + x_2 \le 3, \\ & -x_1 + 2x_2 \le 1; \end{aligned}$$

(g)
$$\begin{aligned} \min\ -4x_1 - 5x_2&,\\ x_1 + 2x_2 &\le 3,\\ x_1 - x_2 &\le 2,\\ 2x_1 + x_2 &\le 3,\\ 3x_1 + 4x_2 &\le 8,\\ x_1 &\ge 0; \end{aligned}$$

(h)
$$\begin{aligned} \max\ 3x_1 + 2x_2 + 4x_3&,\\ x_1 + x_2 + 2x_3 &\le 4,\\ 2x_1 + 3x_3 &\le 5,\\ 2x_1 + x_2 + 3x_3 &\le 7,\\ x_1, x_2, x_3 &\ge 0; \end{aligned}$$

(i)
$$\begin{aligned} \max\ 5x_1 + 6x_2 + 9x_3 + 8x_4&,\\ x_1 + 2x_2 + 3x_3 + x_4 &\le 5,\\ x_1 + x_2 + 2x_3 + 3x_4 &\le 3,\\ x_1, x_2, x_3, x_4 &\ge 0. \end{aligned}$$

3.15. Xét quy hoạch

$$\begin{aligned} \max\ 3x_1 + x_2&,\\ x_1 + 2x_2 &\le 5,\\ x_1 + x_2 - x_3 &\le 2,\\ 7x_1 + 3x_2 - 5x_3 &\le 20,\\ x_1, x_2, x_3 &\ge 0. \end{aligned}$$

Chứng minh rằng nghiệm tối ưu của nó là suy biến.

3.16. Xét quy hoạch

$$\begin{aligned} \max\ x_1 + 2x_2 + 3x_3&,\\ x_1 + x_2 &\le 5,\\ x_1 &\le 1,\\ x_1, x_2, x_3 &\ge 0. \end{aligned}$$

Tìm ít nhất ba nghiệm cơ sở tối ưu và viết biểu thức cho mọi nghiệm không cơ sở tối ưu lập được từ ba nghiệm này.

CHƯƠNG 4

NGHIÊN CỨU HÌNH HỌC QUY HOẠCH TUYẾN TÍNH

4.1 ĐIỂM CỰC BIÊN VÀ NGHIỆM CƠ SỞ

Ở Mục 1.3 ta đã giải quy hoạch tuyến tính hai biến bằng hình học và có nhận xét quan trọng, mà theo cảm nhận trực quan là đúng cả cho trường hợp n biến tổng quát, là nếu quy hoạch tuyến tính có nghiệm tối ưu và miền chấp nhận được (là tập lồi đa diện) có đỉnh thì quy hoạch tuyến tính luôn có nghiệm tối ưu là đỉnh. Chính tính chất này là cơ sở của phương pháp đơn hình. Nội dung của phương pháp là xuất phát từ một đỉnh của miền chấp nhận được; ở mỗi bước lặp ta đi sang một đỉnh kề nó có giá trị hàm mục tiêu nhỏ hơn. Vì số đỉnh là hữu hạn, sau hữu hạn bước ta sẽ đến đỉnh tối ưu.

Tuy vậy, trong hai chương vừa qua ta lại trình bày phương pháp đơn hình bằng lí luận giải tích, dựa vào biểu diễn từ vựng để lập luận đơn giản hơn. Trong chương này ta sẽ xem xét vấn đề từ hình học và sẽ thấy rõ sự tương đương của các cách tiếp cận.

Ta đã định nghĩa tập lồi, tập lồi đa diện và tổ hợp lồi (xem Mục 1.3).

Bao lồi (convex hull) của một tập là tập lồi nhỏ nhất chứa nó. Ta kí hiệu bao lồi của T là coT.

HỆ QUẢ 4.1.1.

(a) Giao của các tập lồi là tập lồi.

(b) Tập lồi đa diện là tập lồi.

(c) Bao lồi của một tập là giao của mọi tập lồi chứa nó.

(d) Bao lồi của một tập là tập mọi tổ hợp lồi của hữu hạn điểm của tập đó.

CHỨNG MINH. (a) Giả sử x và y thuộc $\bigcap_{i \in I} S_i$ và S_i, $i \in I$, là các tập lồi. Với bất kì $\gamma \in [0,1]$ thì $(1-\gamma)x + \gamma y \in S_i$, $i \in I$. Do đó $(1-\gamma)x + \gamma y \in \bigcap_{i \in I} S_i$, tức giao các tập S_i là tập lồi.

(b) Xét tập lồi đa diện $P = \{x \in R^n: Ax \le b\}$, ở đây A là ma trận cấp mXn. Nếu x, y $\in$ P và $\gamma \in [0,1]$ thì $A[(1-\gamma)x + \gamma y] = (1-\gamma)Ax + \gamma Ay \le (1-g)b + \gamma b =$ b tức là $(1-\gamma)x + \gamma y \in P$.

(c) Vì giao của mọi tập lồi chứa tập T, kí hiệu là G, là tập lồi và chứa T nên $G \supset coT$. Mặt khác coT cũng là một tập lồi chứa T nên $G = G \cap coT$. Vậy $G = coT$.

(d) Tập V mọi tổ hợp lồi của hữu hạn điểm của T là tập lồi. V chứa T vì mỗi $x \in T$ là tổ hợp lồi của một điểm nên $x \in V$. Do đó $V \supset coT$. Mặt khác, vì $T \subset coT$ nên V là tập các tổ hợp lồi hữu hạn điểm của coT, tức là $V \subset coT$.□

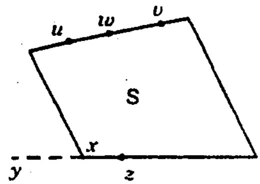
ĐỊNH NGHĨA 4.1.1.

(a) Điểm x của tập lồi S được gọi là *điểm cực biên* (extreme point) của S nếu x không thể là tổ hợp lồi của hai điểm của S khác x, tức là không có y, z $\in$ S, khác x và $\gamma \in [0,1]$ để $x = (1-\gamma)y + \gamma z$.

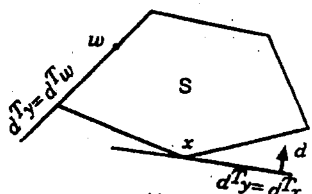
(b) Điểm x của tập lồi S được gọi là *đỉnh* (vertex) của S, nếu có một siêu phẳng sao cho S nằm hoàn toàn ở một phía của nó và siêu phẳng cắt S chỉ ở điểm x này, tức là $\exists d \in R^n$, $d \ne 0$ sao cho $d^Tx < d^Ty$ $\forall y \in S$, $y \ne x$.

Định nghĩa này được minh họa hình học trên H.4.1 và H.4.2 trang bên và cho thấy là khá tự nhiên về trực quan.

Đồng thời ta cũng thấy điểm cực biên và đỉnh khá giống nhau. Dưới đây ta sẽ chứng minh hai khái niệm hoàn toàn trùng nhau, và là điểm biên đặc biệt, "biên nhất có thể được", "biên cùng cực". Tuy nhiên có những tập mà mọi điểm biên đều là điểm cực biên, đều là đỉnh, chẳng hạn như hình cầu đóng trong $\mathbf{R}^n$.



H. 4.1



H. 4.2

H. 4.1 : x là điểm cực biên
w không phải là điểm cực biên
H. 4.2 : x là đỉnh
w không phải là đỉnh

Khi tiến hành thuật toán đơn hình ta đã đi từ một từ vựng sang một từ vựng khác ở mỗi bước. Tương ứng ta đã đi từ nghiệm cơ sở này sang nghiệm cơ sở khác. Ta đã không xét trong thuật toán các nghiệm cơ sở ở mỗi bước đó có phải đỉnh hay điểm cực biên của miền chấp nhận được hay không, vì trong ngôn ngữ thuật toán, khó kiểm tra tính chất hình học này. Bây giờ ta sẽ chứng minh các nghiệm cơ sở ở đây chính là các đỉnh. Ta định nghĩa cụ thể nghiệm cơ sở của tập lồi đa diện, độc lập với ngôn ngữ từ vựng như sau.

ĐỊNH NGHĨA 4.1.2. Xét tập lồi đa diện P trong $\mathbf{R}^n$, xác định bởi các đẳng thức và bất đẳng thức (tuyến tính).

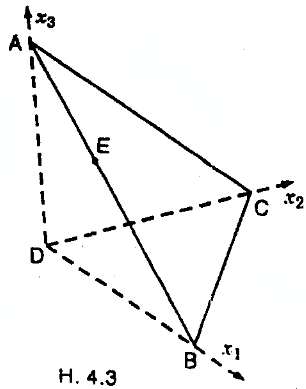
(a) $x^* \in \mathbf{R}^n$ là *nghiệm cơ sở* (basic solution) của P hoặc tương đương là của hệ đẳng thức tuyến tính và bất đẳng thức tuyến tính, nếu nó thỏa mọi ràng buộc đẳng thức và trong số mọi ràng buộc được thỏa chặt, tức là với dấu =, có n ràng buộc độc lập tuyến tính.

(b) Nghiệm cơ sở thỏa mọi ràng buộc được gọi là *nghiệm cơ sở chấp nhận được* (basic feasible solution).

THÍ DỤ. (i) Nghiệm cơ sở ứng với một từ vựng là nghiệm cơ sở của hệ ràng buộc của quy hoạch tuyến tính vì nó thỏa $n - m$ ràng buộc $x_i = 0$, $i \in N$, và m ràng buộc chặt khác là $x_j = ..., j \in B$; tất cả các ràng buộc này độc lập tuyến tính.

Nếu mọi $\bar{b}_j \ge 0, j \in B$, tức là nếu từ vựng là chấp nhận được, thì nghiệm cơ sở tương ứng là chấp nhận được.

(ii) Giả sử $P = \{x \in R^3: x_1 + x_2 + x_3 = 1; x_1, x_2, x_3 \ge 0\}$. Khi đó P chỉ có ba nghiệm cơ sở là ba điểm A, B, C trên H.4.3. Đây cũng là ba nghiệm cơ sở chấp nhận được.



H. 4.3

Nếu ta viết lại P ở dạng

$P = \{x \in R^3: x_1 + x_2 + x_3 \le 1;$
$x_1 + x_2 + x_3 \ge 1; x_1, x_2, x_3 \ge 0\}$

thì điểm D là nghiệm cơ sở của P (nhưng không chấp nhận được). Vậy một điểm là nghiệm cơ sở của tập lồi đa diện hay không, là có phụ thuộc vào dạng của hệ phương trình và bất phương trình biểu diễn P. Mặc dù các dạng khác nhau của cùng tập P phải tương đương.

ĐỊNH LÍ 4.1.2. *Giả sử P là một tập lồi đa diện và $x^* \in P$. Khi đó ta có ba khẳng định tương đương*

(a) x^* là đỉnh;

(b) x^* là điểm cực biên;

(c) x^* là nghiệm cơ sở chấp nhận được.

CHỨNG MINH. (a) $\Rightarrow$ (b). Giả sử x^* là đỉnh, tức là có $d \ne 0$ để $d^T x^* < d^T y$ $\forall y \in P, y \ne x^*$. Với hai điểm y và z khác x^* bất kì của P và với $\gamma \in [0,1]$ bất kì thì $d^T x^* < d^T y$, $d^T x^* \le d^T z$. Do đó $d^T x^* < d^T(\gamma y + (1 - \gamma)z)$. Vì vậy $x^* \ne \gamma y + (1 - \gamma)z$, tức x^* là điểm cực biên.

(b) $\Rightarrow$ (c). Giả sử x^* không phải nghiệm cơ sở chấp nhận được. Ta sẽ chứng tỏ x^* không phải điểm cực biên của P. Giả sử P có biểu diễn là

$\{x \in \mathbf{R}^n: a_i^T x \geq b_i; i = 1, ..., m\}$. Gọi $I = \{i: a_i^T x^* = b_i\}$. Vì x^* không phải là nghiệm cơ sở chấp nhận được, nên trong các vectơ a_i, $i \in I$, không đủ n vectơ độc lập tuyến tính. Do đó tồn tại $d \in \mathbf{R}^n$, $d \neq 0$ sao cho $a_i^T d = 0$, $\forall i \in I$. Xét $y = x^* + \varepsilon d$ và $z = x^* - \varepsilon d$. Ta thấy $a_i^T y = a_i^T x^* = b_i$, $\forall i \in I$. Còn với $i \notin I$ thì $a_i^T x^* > b_i$, nên với ε đủ nhỏ $a_i^T y > b_i$. Vậy $y \in P$ với $\varepsilon > 0$ đủ nhỏ. Tương tự $z \in P$ với $\varepsilon > 0$ đủ nhỏ. Mà $x^* = \frac{1}{2}(y + z)$ nên x^* không phải là điểm cực biên.

(c) $\Rightarrow$ (a). Giả sử $P = \{ x \in \mathbf{R}^n: a_i^T x \geq b_i, i = 1, ..., m \}$ và x^* là nghiệm cơ sở chấp nhận được. Gọi $I = \{ i: a_i^T x^* = b_i \}$ và $d = \sum_{i \in I} a_i$. Khi đó

$$d^T x^* = \sum_{i \in I} a_i^T x^* = \sum_{i \in I} b_i .$$

Với bất kỳ $x \in P$ ta có

$$d^T x = \sum_{i \in I} a_i^T x \geq \sum_{i \in I} b_i . \qquad (4.1)$$

Vì có n vectơ độc lập tuyến tính trong các vectơ a_i, $i \in I$, nên hệ $a_i^T x = b_i$, $i \in I$, có duy nhất nghiệm là x^*. Do đó, với $x \in P$, $x \neq x^*$ ta có $d^T x > d^T x^*$ (do (4.1)). Vậy x^* là đỉnh. □

Do định lí trên, nghiệm cơ sở chấp nhận được của P, cũng như đỉnh và điểm cực biên, không phụ thuộc vào dạng cụ thể của hệ đẳng thức và bất đẳng thức biểu diễn P (khác với nghiệm cơ sở).

MỆNH ĐỀ 4.1.3. *Mỗi hệ ràng buộc tuyến tính chỉ có hữu hạn nghiệm cơ sở. Do đó nó cũng chỉ có hữu hạn nghiệm cơ sở chấp nhận được.*

CHỨNG MINH. Mỗi nghiệm cơ sở thỏa chặt n ràng buộc độc lập tuyến tính.

Vì hệ ràng buộc chặt này có nghiệm duy nhất, nên nghiệm cơ sở khác nhau ứng với hệ ràng buộc chặt độc lập tuyến tính khác nhau. Vì chỉ có số hữu hạn cách chọn n ràng buộc từ số m ràng buộc của hệ, nên chỉ có hữu hạn nghiệm cơ sở. □

4.2 TRƯỜNG HỢP TẬP RÀNG BUỘC CÓ DẠNG CHÍNH TẮC

Bây giờ ta xét trường hợp riêng là P có dạng chính tắc

$$P = \{ x \in \mathbf{R}^n : Ax = b, x \geq 0 \},$$

ở đây A là ma trận cấp $m \times n$. Không mất tổng quát ta giả sử rank $A = m$ (do đó $m \leq n$). Vì nếu có hàng của A phụ thuộc tuyến tính vào các hàng khác và $P \neq \varnothing$ thì ta có thể bỏ hàng đó đi mà vẫn được hệ ràng buộc tương đương.

Ta sẽ thấy định nghĩa nghiệm cơ sở ở Mục 2.1 (qua từ vựng) chính là trường hợp riêng, khi P có dạng chính tắc, của định nghĩa tổng quát ở Mục 4.1.

ĐỊNH LÍ 4.2.1 *(nghiệm cơ sở). Vecto $x \in \mathbf{R}^n$ là nghiệm cơ sở của tập lồi đa diện P ở dạng chính tắc khi và chỉ khi $Ax = b$ và có m chỉ số $B(1), \ldots, B(m)$ sao cho*

(a) các cột $A_{B(1)}, \ldots, A_{B(m)}$ của A là độc lập tuyến tính;

(b) nếu $i \notin \{B(1), \ldots, B(m)\}$ thì $x_i = 0$.

CHỨNG MINH. *Khi.* Do (a) và (b) ta có

$$\sum_{i=1}^{m} A_{B(i)} x_{B(i)} = \sum_{j=1}^{n} A_j x_j = Ax = b.$$

Xét hệ m ẩn $x_{B(1)}, \ldots, x_{B(m)}$, ta thấy các cột $A_{B(1)}, \ldots, A_{B(m)}$ độc lập tuyến tính, nên nghiệm $x_{B(1)}, \ldots, x_{B(m)}$ xác định duy nhất. Bổ sung thêm các thành phần $x_i = 0$, $i \notin \{B(1), \ldots, B(m)\}$ ta được nghiệm xác định duy nhất của hệ n ẩn

$$Ax = b,$$

$$x_i = 0,\ i \notin \{B(1), ..., B(m)\}.$$

Do đó n phương trình phải độc lập tuyến tính. Vậy x thỏa chặt n ràng buộc độc lập tuyến tính, tức x là nghiệm cơ sở.

Chỉ khi. Giả sử x là nghiệm cơ sở, có các thành phần khác 0 là $x_{B(1)}, ..., x_{B(k)}$. Vì x thoả n ràng buộc chặt độc lập tuyến tính

$$Ax = b,$$

$$x_i = 0,\ i \notin \{B(1), ..., B(k)\},$$

nên hệ này có duy nhất nghiệm. Thay $x_i = 0,\ i \notin \{B(1), ..., B(k)\}$ vào phương trình đầu ta được $\sum_{i=1}^{k} A_{B(i)} x_{B(i)} = b$, nên hệ này cũng có duy nhất nghiệm. Do đó $A_{B(1)}, ..., A_{B(k)}$ độc lập tuyến tính. Vậy $k \le m$. Vì A có m cột độc lập tuyến tính nên khi $k < m$ ta có thể bổ sung thêm $A_{B(k+1)}, ..., A_{B(m)}$ để được m cột độc lập tuyến tính. Rõ ràng $x_i = 0$ với $i \notin \{B(1), ..., B(m)\}$.□

Theo Định lí 4.2.1 để có một nghiệm cơ sở ta chỉ việc chọn m cột độc lập tuyến tính $A_{B(1)}, ..., A_{B(m)}$ trong A và đặt $x_i = 0$ khi $i \notin \{B(1), ..., B(m)\}$, rồi giải hệ m phương trình $Ax = b$ ra các thành phần còn lại $x_{B(1)}, ..., x_{B(m)}$. Đây chính là định nghĩa nghiệm cơ sở ở Mục 2.1 và Mục 3.3. Khi đó $A_{B(1)}, ..., A_{B(m)}$ chính là cơ sở của nghiệm cơ sở x theo Định nghĩa 3.3.1. Thông thường, khi không sợ nhầm lẫn người ta cũng lấy ngay tập chỉ số B của cơ sở $\{A_i, i \in B\}$ để kí hiệu cơ sở đó. Khi đó ta có $Bx_B = b$ và $x_B = B^{-1}b$ (kí hiệu x_B là để chỉ rằng đây là nghiệm cơ sở ứng với cơ sở B).

Theo định nghĩa cơ sở, ứng với một cơ sở chỉ có một nghiệm cơ sở duy nhất. Ngược lại, một nghiệm cơ sở có thể ứng với nhiều cơ sở khác nhau. Một thí dụ đặc biệt là khi $b = 0$ thì mọi cơ sở đều ứng với một nghiệm cơ sở là vectơ 0. Trường hợp một nghiệm cơ sở ứng với nhiều cơ sở khác nhau liên quan mật thiết với hiện tượng suy biến của nghiệm cơ sở như ta đã nói đến ở Mục 3.3 khi định nghĩa cơ sở, và sẽ thấy rõ hơn ở Mục 4.4 dưới đây.

4.3 SỰ TỒN TẠI ĐỈNH VÀ ĐỈNH TỐI ƯU

Ta đã biết rằng, đối với tập lồi đa diện bất kì, ba khái niệm đỉnh, điểm cực biên và nghiệm cơ sở chấp nhận được là trùng nhau. Nghiệm cơ sở chấp nhận được là then chốt của phương pháp đơn hình. Ở Mục 2.1 ta đã biết là tập lồi đa diện có thể không có đỉnh. Bây giờ ta xét điều kiện để tồn tại đỉnh.

Chú ý rằng một đường thẳng trong $\mathbf{R}^n$ qua điểm $x \in \mathbf{R}^n$ có thể biểu diễn ở dạng $\{ x + \lambda d: \lambda \in \mathbf{R} \}$, ở đây $d \neq 0$, $d \in \mathbf{R}^n$ chính là vectơ chỉ phương của đường thẳng (H.4.4).

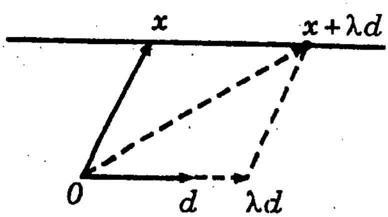
ĐỊNH LÍ 4.3.1. *Giả sử tập lồi đa diện $P = \{ x: a_i^T x \geq b_i,\ i = 1, ..., m \}$ khác trống. Khi đó các kết luận sau là tương đương:*

(a) P có đỉnh;

(b) P không chứa một đường thẳng nào;

(c) trong các vectơ $a_1, ..., a_m$ có n vectơ độc lập tuyến tính.

CHỨNG MINH: (a) $\Rightarrow$ (c). Nếu x là đỉnh thì nó thỏa n ràng buộc chặt độc lập tuyến tính trong số $a_i^T x \geq b_i$, $i = 1, ..., m$. Khi đó n vectơ a_i tương ứng là độc lập tuyến tính.



H. 4.4

(c) $\Rightarrow$ (b). Không mất tổng quát giả sử n vectơ độc lập tuyến tính là $a_1, ..., a_n$. Giả sử phản chứng là P chứa đường thẳng $\{ x + \lambda d: \lambda \in R\}$, ở đây $d \neq 0$. Khi đó $a_i^T(x + \lambda d) \geq b_i$, $i = 1, ..., m$, với mọi λ.

Từ đây ta thấy $a_i^T d = 0$, $i = 1, ..., m$, (vì giả sử có $a_k^T d > 0$ thì lấy $\lambda \to -\infty$ ta có $a_k^T(x + \lambda d) \to -\infty$ mâu thuẫn với bất đẳng thức $a_k^T(x + \lambda d) \geq b_k$ với mọi λ; tương tự $a_k^T d < 0$ không thể xảy ra). Nhưng $a_1, \dots, a_n$ độc lập tuyến tính nên từ $a_i^T = 0$, $i = 1, ..., n$, suy ra $d = 0$, trái với định nghĩa d.

(b) ⇒ (a). Giả sử $x \in P$ và $I := \{ i: a_i^T x = b_i \}$. Nếu có n vectơ a_i, $i \in I$, độc lập tuyến tính, thì x là nghiệm cơ sở chấp nhận được, tức là điểm cực biên rồi.

Trường hợp ngược lại thì mọi vectơ a_i, $i \in I$, nằm trong một không gian con thực sự của $\mathbf{R}^n$. Do đó sẽ có $d \in \mathbf{R}^n$, $d \neq 0$, vuông góc với chúng, tức là $a_i^T d = 0$, $i \in I$. Xét đường thẳng $\{x + \lambda d: \lambda \in \mathbf{R}\}$. Ta có

$$a_i^T(x + \lambda d) = a_i^T x + \lambda a_i^T d = a_i^T x = b_i, \ \forall i \in I, \ \forall \lambda \in \mathbf{R}.$$

tức là các ràng buộc chặt tại x vẫn thỏa chặt trên cả đường thẳng. Nhưng P không chứa đường thẳng nên khi thay đổi λ để đi theo đường thẳng phải đến một giá trị λ^* để qua đó có ràng buộc j (chỉ số j không thuộc I) bị vi phạm. Vậy điểm ranh giới phải thỏa chặt ràng buộc j, tức là $a_j^T(x + \lambda^* d) = b_j$.

Ta sẽ chỉ rằng a_j không phải tổ hợp tuyến tính của các a_i, $i \in I$. Ta có $a_j^T x \neq b_j$ (vì $j \notin I$). Do đó $a_j^T d \neq 0$, tức là d không vuông góc với a_j mà lại vuông góc với mọi a_i, $i \in I$. Vậy a_j không phải tổ hợp tuyến tính của các a_i, $i \in I$. Như vậy đi từ x đến $x + \lambda^* d$, số các ràng buộc chặt độc lập tuyến tính đã thêm ít nhất là một. Cứ tiếp tục quá trình, lại đi từ điểm $x + \lambda^* d$, ... cuối cùng ta phải đi đến điểm thỏa chặt n ràng buộc độc lập tuyến tính, chính là một đỉnh. □

HỆ QUẢ 4.3.2. *Tập lồi đa diện ở dạng chính tắc mà khác trống thì luôn có đỉnh.*

CHỨNG MINH. $P = \{ x: Ax = b, x \geq 0 \}$ nằm trong góc tọa độ dương $\{x: x \geq 0\}$ nên không thể chứa đường thẳng. □

NHẬN XÉT 4.3.1 Quy hoạch tuyến tính có thể có miền ràng buộc không có đỉnh. Để tiến hành thuật toán đơn hình ta đưa nó về dạng chính tắc bằng cách thêm biến bù. Khi đó theo Hệ quả 4.3.2, trong không gian các biến mới (nhiều chiều hơn) tập ràng buộc luôn có đỉnh. Đây là một lí do để ta phải đưa về dạng chính tắc khi tiến hành thuật toán đơn hình. Ở Mục 2.3, khi đưa ra bài toán bổ trợ để tìm từ vựng xuất phát, ta đã thừa nhận là luôn có từ vựng này. Hệ quả trên mới chứng minh điều thừa nhận này. Người ta cũng còn nhiều phương pháp khác để tìm từ vựng xuất phát mà chúng tôi sẽ đề cập đến, cùng với nhiều nội dung khác, trong sách "Quy hoạch tuyến tính. Kiến thức bổ sung và bài tập".

Dạng chính tắc của tập lồi đa diện còn thuận lợi ở một điểm nữa là đỉnh của nó, tức là nghiệm cơ sở, có dạng đơn giản như Định lí 4.2.1 khẳng định. Chính điều đó cho phép ta tiến hành thuật toán đơn hình.

ĐỊNH LÍ 4.3.3 (tồn tại đỉnh tối ưu). *Xét bài toán min c^Tx trên tập lồi đa diện $P = \{ x: Ax \geq b \}$ khác trống. Giả sử P có đỉnh. Khi đó hoặc hàm mục tiêu không giới nội dưới trên P hoặc P có đỉnh tối ưu.*

CHỨNG MINH. Ta gần như lặp lại chứng minh Định lí 4.3.1, chỉ có điểm khác là khi đi dần về một đỉnh của P, ta làm sao để có thêm tính chất là hàm mục tiêu không tăng.

Ta sẽ nói $x \in P$ là có hạng bằng k nếu ta chỉ tìm được đến k ràng buộc độc lập tuyến tính thỏa chặt tại x. Ta chỉ phải xét trường hợp hàm mục tiêu giới nội dưới trên P. Lấy $x \in P$ bất kì và giả sử hạng của x là $k < n$.

Trước hết ta chứng tỏ sẽ có $y \in P$ có hạng lớn hơn x và $c^Ty \leq c^Tx$. Đặt $I = \{i: a_i^Tx = b_i\}$. Vì $k < n$, các vectơ a_i nằm trong một không gian con thực sự của $\mathbf{R}^n$ và sẽ có $d \neq 0$ để $a_i^Td = 0$, $i \in I$. Không mất tổng quát, giả sử $c^Td \leq 0$ (nếu $c^Td > 0$ thì lấy $-d$ thay cho d).

Nếu $c^Td < 0$, ta xét nửa đường thẳng $\{ y = x + \lambda d: \lambda > 0 \}$. Ta có $a_i^T(x + \lambda d) = a_i^Tx = b_i$, $i \in I$, $\lambda > 0$, nên các ràng buộc thỏa chặt ở x cũng thỏa chặt trên mọi điểm của nửa đường thẳng. Vì $c^T(x + \lambda d) \to -\infty$ khi $\lambda \to +\infty$ nên đến giá trị λ^* phải có ràng buộc $j \notin I$ nào đó là ràng buộc chặt $a_j^T(x + \lambda^* d) = b_j$ để sau đó $x + \lambda d$ ra ngoài P. Cũng như ở Định lí 4.3.1 ta thấy hạng của $y := x + \lambda^* d$ ít nhất là $k + 1$ vì a_j độc lập tuyến tính với các a_i, $i \in I$. Hơn nữa ta cũng có $c^T y < c^T x$ (vì $c^T d < 0$).

Nếu $c^T d = 0$, ta xét cả đường thẳng $\{ y = x + \lambda d : \lambda \in R \}$ và tương tự trên ta sẽ thấy điểm $x + \lambda^* d$, khi đường thẳng đến ranh giới để rời P, có hạng lớn hơn hạng của x. Hơn nữa $c^T(x + \lambda^* d) = c^T x$.

Cứ tiếp tục quá trình cuối cùng ta sẽ đến một điểm w có hạng n, tức w là một đỉnh của P, mà $c^T w \le c^T x$.

Giả sử $w^1, \ldots, w^r$ là các đỉnh của P (vì chỉ có hữu hạn đỉnh) và w^* là đỉnh có $c^T w^* \le c^T w^i$, $i = 1, \ldots, r$. Vì ta đã chỉ ra là với mọi $x \in P$, sẽ có một w^i để $c^T w^i \le c^T x$, nên $c^T w^* \le c^T x$, $\forall x \in P$. Vậy w^* là đỉnh tối ưu. □

HỆ QUẢ 4.3.4. *Bài toán quy hoạch tuyến tính min $c^T x$ trên tập lồi đa diện khác trống hoặc không giới nội dưới hoặc là có nghiệm tối ưu.*

CHỨNG MINH. Ta đưa bài toán về dạng chính tắc tương đương với nó. Khi đó miền ràng buộc luôn có đỉnh và áp dụng được Định lí 4.3.3. Tuy nhiên đỉnh tối ưu theo Định lí 4.3.3 lại chỉ cho ta nghiệm tối ưu của tập lồi đa diện ban đầu, chưa chắc là đỉnh. □

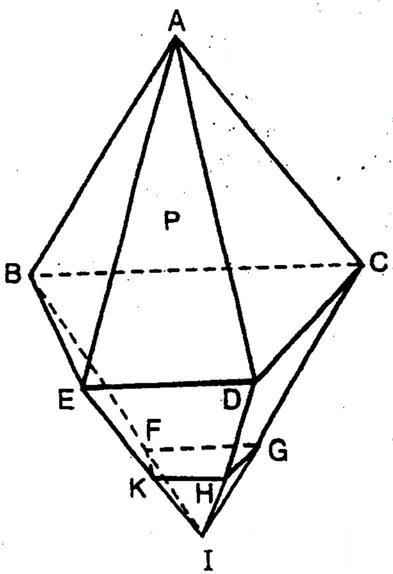
Sự tồn tại nghiệm tối ưu như trên là đặc thù của quy hoạch tuyến tính. Bài toán phi tuyến thậm chí đơn giản, như min $\{ \frac{1}{x} : x \ge 1 \}$ cũng không có nghiệm tối ưu mặc dù hàm mục tiêu ở đây lồi, liên tục, giới nội dưới và miền ràng buộc là tập lồi đa diện một chiều.

4.4 HIỆN TƯỢNG SUY BIẾN

ĐỊNH NGHĨA 4.4.1. Nghiệm cơ sở của hệ ràng buộc tuyến tính được gọi là *suy biến* nếu nó thỏa chặt hơn n ràng buộc.

Như vậy tại nghiệm cơ sở đã phải có n ràng buộc độc lập tuyến tính thỏa chặt, lại có thêm ràng buộc thỏa chặt hơn mức cần thiết thì nghiệm cơ sở sẽ là suy biến (tất nhiên trong $\mathbf{R}^n$ chỉ có đến n ràng buộc thỏa chặt độc lập tuyến tính).

THÍ DỤ. (i) Nếu hệ ràng buộc tuyến tính viết ở dạng từ vựng thì nghiệm cơ sở tương ứng là suy biến nếu có hơn n − m thành phần bằng 0, tức là có $b_i = 0$. Vậy định nghĩa nghiệm cơ sở suy biến (ứng với từ vựng suy biến) ở Mục 3.1 hoàn toàn phù hợp và là trường hợp riêng của Định nghĩa 4.4.1. Định nghĩa nghiệm cơ sở suy biến của từ vựng thực ra là định nghĩa nghiệm cơ sở suy biến của ràng buộc tuyến tính ở dạng chính tắc $Ax = b$, $x \geq 0$, nói chung. Đối với hệ ràng buộc dạng chính tắc ta có khái niệm cơ sở. Vì nghiệm cơ sở suy biến có hơn n − m thành phần bằng 0, nên nói chung ta có nhiều cách chọn n − m biến trong đó làm biến cơ sở. Do đó sẽ có nhiều cơ sở ứng với cùng nghiệm cơ sở suy biến đó. (Đây là trường hợp điển hình. Tuy nhiên có thí dụ cho thấy nghiệm cơ sở suy biến có thể chỉ có một cơ sở tương ứng duy nhất.)



H. 4.5

(ii) Trong không gian $\mathbf{R}^3$, nghiệm cơ sở suy biến thỏa hơn 3 ràng buộc chặt (tuyến tính), có nghĩa là nó là giao của hơn 3 mặt phẳng. Chẳng hạn ở tập lồi đa diện trong hình H.4.5, đỉnh A, B, C, D và E là nghiệm cơ sở suy biến chấp nhận được (cũng có thể gọi là *đỉnh suy biến*). F, G, H và K là nghiệm cơ sở không suy biến, chấp nhận được. I là nghiệm cơ sở suy biến (không chấp nhận được).

CHÚ Ý. Hiện tượng suy biến không phải là một tính chất thuần túy hình học, độc lập với dạng biểu diễn giải tích. Trái lại nó phụ thuộc vào dạng biểu diễn bởi các bất đẳng thức và đẳng thức tuyến tính.

THÍ DỤ. (i) Xét tập lồi đa diện ở dạng chính tắc

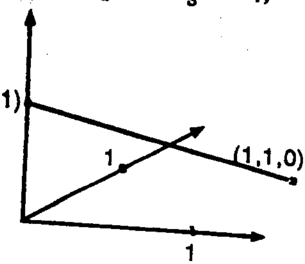
$$P = \{(x_1, x_2, x_3): x_1 - x_2 = 0; x_1 + x_2 + 2x_3 = 2; x_1, x_2, x_3 \geq 0 \}.$$

Vì n = 3, m = 2, n − m = 1, nên (1,1,0) là đỉnh không suy biến còn (0,0,1) là đỉnh suy biến. P là đoạn thẳng nối hai đỉnh này (H.4.6).

Nếu ta biểu diễn lại P ở dạng khác (không phải chính tắc):

$$P = \{ (x_1, x_2, x_3): x_1 - x_2 = 0; x_1 + x_2 + 2x_3 = 2; x_1 \geq 0, x_3 \geq 0 \},$$

thì (0,0,1) là đỉnh không suy biến vì nó chỉ thỏa chặt đúng ba ràng buộc (độc lập tuyến tính).



H. 4.6

(ii) Thậm chí, việc chuyển dạng chính tắc tổng quát sang dạng chuẩn tổng quát bao giờ cũng làm thay đổi tính không suy biến. Giả sử x* là đỉnh không suy biến của $P = \{ x: Ax = b, x \geq 0 \}$, ở đây A là ma trận cấp m x n có hạng là m. Khi đó có đúng n − m thành phần $x_i^* = 0$. Bây giờ xét P ở dạng chuẩn, $P = \{x: Ax \leq b, -Ax \leq -b, x \geq 0\}$. Tất nhiên x* vẫn là nghiệm cơ sở chấp nhận được, tức là đỉnh. Nhưng còn tính suy biến thì sao? Tại x* có n − m ràng buộc dấu thỏa chặt và 2m ràng buộc bất đẳng thức thỏa chặt. Vậy có m + n ràng buộc thỏa chặt nên x^* là đỉnh suy biến của P ở dạng biểu diễn chuẩn.

Tuy nhiên, độc giả hãy chứng minh, coi như bài tập, rằng một nghiệm cơ sở chấp nhận được là suy biến ở một dạng chính tắc của P thì vẫn phải là suy biến ở mọi dạng chính tắc khác của P.

4.5 GIẢI THÍCH HÌNH HỌC PHƯƠNG PHÁP ĐƠN HÌNH

Ta đã biết một số sự kiện hình học sau đây

- Khi thực hiện thuật toán đơn hình ta biểu diễn tập ràng buộc ở dạng chính tắc. Về hình học khi đó không gian các biến là thích hợp để tập ràng buộc luôn có đỉnh.

- Ở mỗi bước lặp của thuật toán đơn hình ta chuyển từ một nghiệm cơ sở chấp nhận được sang một nghiệm cơ sở chấp nhận được khác có giá trị mục tiêu tốt hơn. Về hình học có nghĩa là chuyển từ một đỉnh của tập ràng buộc sang đỉnh khác có giá trị mục tiêu tốt hơn.

Nhưng vị trí hai đỉnh đó so với nhau như thế nào? Bây giờ ta xem xét.

Tập các điểm cùng thỏa chặt một ràng buộc tuyến tính, tức là $\{x: a_i^T x = b_i\}$ là một siêu phẳng trong $\mathbf{R}^n$, tức là một diện n − 1 chiều. Tập các điểm cùng thỏa chặt hai ràng buộc độc lập tuyến tính là giao của hai siêu phẳng, tức là một diện n − 2 chiều. Cứ tiếp tục như vậy ta thấy tập các điểm cùng thỏa chặt n − 1 ràng buộc độc lập tuyến tính sẽ là một cạnh (diện một chiều) của tập lồi đa diện. Và cuối cùng tập các điểm thỏa chặt n ràng buộc độc lập tuyến tính là một diện 0 chiều, tức là một điểm mà ta đã gọi là đỉnh hoặc điểm cực biên.

Như vậy hai đỉnh của tập lồi đa diện trong $\mathbf{R}^n$ sẽ gọi là *kề nhau* (adjacent) nếu có n − 1 ràng buộc độc lập tuyến tính thỏa chặt tại cả hai đỉnh.

Bây giờ xét hai đỉnh ở hai bước lặp liền nhau của thuật toán đơn hình. Khi chuyển sang nhau, chỉ có một biến x_k vào cơ sở, tức là một ràng buộc chặt $x_k = 0$ trở thành thỏa không chặt $x_k > 0$. Đồng thời, để bù lại, cũng có đúng một biến cơ sở x_i trở thành không cơ sở, tức là có một ràng buộc thỏa không chặt $x_i > 0$ lại thành thỏa chặt $x_i = 0$. Như vậy chỉ đổi chỉ số một ràng buộc chặt (k thành i), tức là còn n − 1 ràng buộc độc lập tuyến tính thỏa chặt ở cả hai đỉnh. Vậy hai đỉnh này là kề nhau. Ta có kết luận hình học sau đây.

- Ở mỗi bước lặp của thuận toán đơn hình ta đã đi từ một đỉnh, theo một *cạnh* (edge) để đến một đỉnh kề có giá trị mục tiêu tốt hơn.

Tiếp theo, ta giải thích hình học phương pháp nhiễu loạn để tránh xoay vòng. Chẳng hạn, xét đỉnh suy biến A ở H.4.5 trang 102.

A suy biến vì nó thỏa chặt bốn ràng buộc (trong không gian ba chiều). Về hình học, điều này nghĩa là A là giao điểm của bốn mặt phẳng biên của miền chấp nhận được P. Như vậy là quá "mức bình thường", mà đã suy biến, vì chỉ cần ba mặt phẳng không có cặp nào song song nhau là xác định một giao điểm duy nhất. Tuy vậy, sự suy biến này rất ngẫu nhiên, "tình cờ". Chỉ một nhiễu nhỏ, vào một mặt phẳng trong bốn mặt chẳng hạn, đã có thể làm bốn mặt không có một điểm chung nữa, mà chỉ từng ba mặt có điểm chung. Khi đó một đỉnh suy biến A trở thành vài đỉnh không suy biến. Chính vì vậy khi nhiễu cả m ràng buộc với ε_i, $i = 1, ..., m$, khác hẳn cỡ nhau, ta đã loại bỏ được hoàn toàn các đỉnh suy biến, như đã chứng minh bằng giải tích.

4.6 TẠI SAO GỌI LÀ THUẬT TOÁN ĐƠN HÌNH?

Trong mục này ta sẽ đưa ra một cách giải thích hình học khác cho phương pháp đơn hình, cũng thường được gọi là *hình học theo cột* (column geometry) của phương pháp đơn hình. Cách tiếp cận này giải thích được nguồn gốc của thuật ngữ "phương pháp đơn hình" và "phép xoay". Đồng thời nó cũng cho thấy tại sao phương pháp đơn hình lại thường có hiệu quả cao, thường giải khá nhanh các bài toán cỡ lớn.

Khác với trước đây, ta viết quy hoạch tuyến tính ở dạng

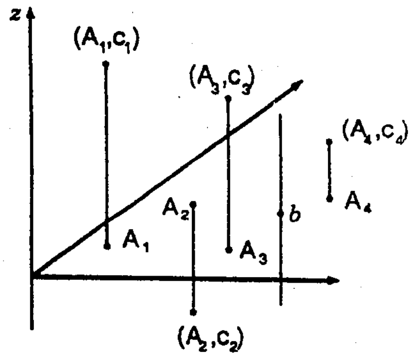
$$
\begin{aligned}
&\min c^T x,\\
&Ax = b,\\
&x_1 + x_2 + \ldots + x_n = 1,\\
&x \geq 0,
\end{aligned}
\qquad (4.2)
$$

ở đây A là m x n ma trận với rank A = m. Ràng buộc $x_1 + x_2 + \ldots + x_n = 1$ được gọi là *ràng buộc lồi* (convexity constraint). Độc giả hãy tự chứng minh như một bài tập rằng mọi quy hoạch tuyến tính có miền chấp nhận được giới nội đều được đưa về dạng (4.2).

Để biểu diễn hình học ta đặt $z = c^T x$, gọi các cột của A là $A_1, \ldots, A_n$ và viết lại (4.2) như sau

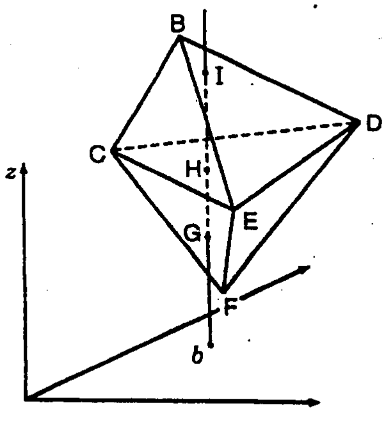
$$\min z,$$
$$x_1 + x_2 + \ldots + x_n = 1,\ x_1, \ldots, x_n \geq 0,$$
$$x_1\begin{pmatrix} A_1 \\ c_1 \end{pmatrix} + x_2\begin{pmatrix} A_2 \\ c_2 \end{pmatrix} + \ldots + x_n\begin{pmatrix} A_n \\ c_n \end{pmatrix} = \begin{pmatrix} b \\ z \end{pmatrix}. \tag{4.3}$$

Trên H.4.7 mặt phẳng ngang biểu diễn không gian m chiều, sinh bởi các cột $A_1, \ldots, A_n$. Trục thẳng đứng là một chiều, biểu diễn giá trị của z và các c_j. Khi đó ta có n điểm $(A_1, c_1), \ldots, (A_n, c_n)$ trong không gian m + 1 chiều (tưởng tượng là ba chiều trên H.4.7!). Các điểm dạng (b, z) với z thay đổi tạo thành đường thẳng đứng, gọi là *đường đòi hỏi* (requirement line). Vậy bài toán (4.3) trở thành bài toán tìm điểm (b, z) trên đường đòi hỏi là tổ hợp lồi của n điểm $(A_1, c_1), \ldots, (A_n, c_n)$ sao cho cao độ z là nhỏ nhất.



H. 4.7

Nếu đường đòi hỏi không cắt bao lồi của các điểm (A_j, c_j), bài toán là không chấp nhận được. Nếu có cắt thì có nghiệm chấp nhận được và nghiệm tối ưu tương ứng với điểm thấp nhất ở giao của đường đòi hỏi với tổ hợp lồi các (A_j, c_j). Độ cao của điểm này là giá trị mục tiêu tối ưu. Chẳng hạn, trong H.4.8 thì nghiệm tối ưu tương ứng với điểm G.



H. 4.8

Sau khi biểu diễn theo hình học cột bài toán quy hoạch tuyến tính, ta xác định rõ ý nghĩa của thuật toán "đơn hình" bởi định nghĩa sau.

ĐỊNH NGHĨA 4.6.1.

(i) Bộ k + 1 điểm $y^1, \ldots, y^{k+1}$ trong R^n (cũng là k + 1 vectơ) được gọi là *độc lập affine* (affinely independent) nếu k vectơ $y^1 - y^{k+1}, \ldots, y^k - y^{k+1}$ là độc lập tuyến tính.

(ii) Bao lồi của k + 1 điểm độc lập affine trong R^n được gọi là *đơn hình k chiều* (k – dimensional simplex).

Nhận xét rằng ở định nghĩa trên ta phải có $k \le n$. Các trường hợp ít chiều nhất, theo định nghĩa trên là như sau. Một điểm là đơn hình 0 chiều. Đoạn thẳng nối hai điểm là đơn hình 1 chiều. Ba điểm thì hoặc là nằm trên một đường thẳng hoặc là độc lập affine và xác định đơn hình hai chiều (tam giác). Bốn điểm hoặc là nằm trên cùng một mặt phẳng hoặc là độc lập affine và xác định đơn hình ba chiều (tứ diện).... Như vậy đơn hình k chiều là hình k chiều "đơn giản nhất" theo định nghĩa là xác định bằng số điểm tối thiểu để được hình lồi (đóng) có phần trong tương đối k chiều.

Bây giờ hãy giải thích hình học nghiệm cơ sở chấp nhận được của quy hoạch tuyến tính (4.2). Vì hệ có m+1 ràng buộc đẳng thức và ta giả sử là độc lập tuyến tính, nên mỗi nghiệm cơ sở chấp nhận được ứng với một cơ sở gồm m+1 cột độc lập tuyến tính $\begin{pmatrix} A_i \\ 1 \end{pmatrix}$, i = 1, ..., m+1, (là các vectơ m+1 chiều). m+1 điểm $\begin{pmatrix} A_i \\ c_i \end{pmatrix}$, i = 1, ..., m+1, tương ứng được gọi là các điểm cơ sở, các $\begin{pmatrix} A_j \\ c_j \end{pmatrix}$ khác được gọi là điểm không cơ sở. Độc giả hãy chứng minh, như một bài tập, rằng m+1 điểm cơ sở là độc lập affine. Bao lồi của nó là đơn hình m chiều và được gọi là *đơn hình cơ sở* (basic simplex). Giả sử đường đòi hỏi cắt đơn hình cơ sở ở điểm $(b,\bar{z})$ nào đó. Khi đó các hệ số $(x_1, ..., x_n)$ của biểu diễn $(b,\bar{z})$ thành tổ hợp lồi của các điểm cơ sở $\begin{pmatrix} A_i \\ c_i \end{pmatrix}$, $i = 1, ..., m+1$, sẽ có $n - (m+1)$ thành phần bằng 0. Do đó $(x_1, ..., x_n)$ là nghiệm cơ sở chấp nhận được của quy hoạch (4.2) và $\bar{z}$ là giá trị mục tiêu tương ứng. Chẳng hạn ở H.4.8 thì tam giác CDE là đơn hình cơ sở và điểm H biểu thị nghiệm cơ sở chấp nhận được liên quan đến các điểm cơ sở C, D và E.

Tiếp theo ta xét xem một bước lặp của thuật toán đơn hình có nghĩa như thế nào trong hình học theo cột. Ta gọi mặt phẳng m chiều đi qua m + 1 điểm cơ sở là mặt phẳng đối ngẫu (dual plane). Ở một bước lặp, một cột trong cơ sở $\begin{pmatrix} A_i \\ 1 \end{pmatrix}$ trở thành không cơ sở và một cột ngoài vào cơ sở. Tương ứng là một điểm cơ sở $\begin{pmatrix} A_i \\ c_i \end{pmatrix}$ bị thay bởi một điểm không cơ sở. Việc thay này theo phương pháp đơn hình là phải đảm bảo để giá trị mục tiêu tương ứng giảm. Về hình học điều này có nghĩa là đường đòi hỏi phải cắt đơn hình cơ sở mới ở một điểm thấp hơn. Điều này xảy ra khi và chỉ khi điểm cơ sở mới vào nằm dưới mặt phẳng đối ngẫu. Ở thí dụ trên hình H.4.8 với đơn hình cơ sở hiện tại là tam giác CDE thì trong hai điểm ngoài cơ sở là B và F, chỉ có F thỏa điều kiện vào.

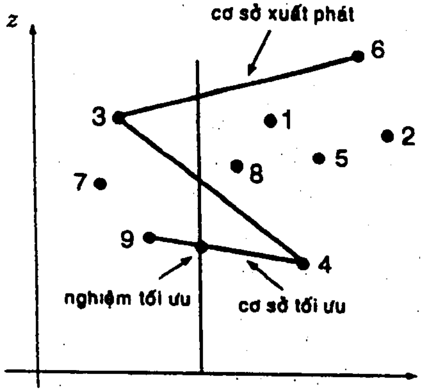
Khoảng cách thẳng đứng từ mặt phẳng cơ sở đến một điểm $\begin{pmatrix} A_j \\ c_j \end{pmatrix}$ chính là bằng hệ số của hàm mục tiêu ứng với biến x_j (độc giả hãy chứng minh!). Vì vậy theo *quy tắc xoay Dantzig* (Dantzig's pivot rule), tức là quy tắc theo đó khi chọn biến vào ta chọn biến có hệ số âm, nhỏ nhất, điểm ngoài cơ sở được chọn vào là điểm ở dưới mặt phẳng đối ngẫu và cách xa nó nhất theo chiều thẳng đứng.

Còn điểm cơ sở sẽ ra khỏi cơ sở được xác định về hình học bằng cách nào? Nhận xét rằng đơn hình cơ sở hiện hữu (lúc đang chuẩn bị một bước lặp) có m + 1 điểm cơ sở và ta đã có thêm một điểm ngoài sẽ vào vừa chọn như trên. Cứ chọn một điểm cơ sở cho ra thì ta còn m điểm cùng với điểm sẽ vào lại lập thành một đơn hình m chiều. Vì có m + 1 cách chọn một điểm ra nên ta có m + 1 đơn hình m chiều tương ứng. Cùng với đơn hình cơ sở hiện hữu chúng tạo thành m + 2 đơn hình m chiều lập thành từ m + 2 điểm. m + 2 đơn hình này chính là m + 2 diện của đơn hình m + 1 chiều do m + 2 điểm đó tạo thành. Chẳng hạn trong H.4.8 (m = 2), sau khi đã có điểm F vào đơn hình cơ sở hiện hữu là tam giác CDE, thì ba cách chọn điểm ra (C,D và E) cùng với F lập thành ba đơn hình hai chiều CDF, CEF và DEF. Cùng với đơn hình cơ sở hiện hữu CDE, chúng lập thành 4 diện của đơn hình m + 1 chiều CDEF.

Đơn hình cơ sở hiện hữu phải cắt đường đòi hỏi (ở H.4.8 là tại H), tức là đường đòi hỏi đi ra khỏi đơn hình m + 1 chiều đang nói đến (CDEF ở H.4.8). Do đó nó phải vào đơn hình m + 1 chiều này ở một diện phía dưới (ở H.4.8 là diện CEF). Diện này là đơn hình m chiều gồm m + 1 điểm; trong đó có m điểm của đơn hình cơ sở hiện hữu. Còn lại một điểm khác của đơn hình cơ sở này chính là điểm sẽ ra khỏi cơ sở (ở H.4.8 là điểm D).

Như vậy sau khi xác định xong một điểm cơ sở ra và một điểm ngoài vào ta được một đơn hình cơ sở mới (ở H.4.8 là CEF). Ta hãy tưởng tượng bước lặp này là gỡ một đỉnh của đơn hình cơ sở (đỉnh D) còn các đỉnh khác vẫn gắn chặt, rồi xoay đơn hình này cho đỉnh đã gỡ ra rời đến đỉnh mới (là điểm vào F). Như vậy là ta đã có phép xoay đơn hình cơ sở ở mỗi bước. Đây chính là nguồn gốc của thuật ngữ "phép xoay" và "phương pháp đơn hình".

THÍ DỤ 4.6.1. Để hình dung hiệu quả của phương pháp đơn hình ta xét bài toán với m = 1, n = 10 được minh họa hình học theo cột ở H.4.9, ở đây số i là



H. 4.9

tương ứng với điểm (A_i, C_i). Giả sử đơn hình cơ sở xuất phát gồm hai điểm (A_3, C_3) và (A_6, C_6), tức là đoạn nối 3 và 6 và ta dùng quy tắc xoay Dantzig. Chỉ sau hai bước xoay đơn hình cơ sở, đến đoạn 3-4 rồi đến đoạn 4-9 ta đã được cơ sở tối ưu.

BÀI TẬP CHƯƠNG 4

4.1. Chứng minh rằng một nghiệm cơ sở chấp nhận được là suy biến ở một dạng chính tắc của tập lồi đa diện P thì vẫn là suy biến ở mọi dạng chính tắc khác của P. (Nhớ rằng hiện tượng suy biến phụ thuộc vào dạng biểu diễn tập lồi đa diện bởi các đẳng thức và bất đẳng thức.)

4.2. Đường thẳng qua hai điểm a, b $\in R^n$ là tập

$$\{x \in R^n: x = (1 - \lambda)a + \lambda b, \lambda \in R\}$$

(hãy biểu diễn hình học!) Hãy viết định nghĩa ở dạng tập tương tự cho đoạn [a, b] trên đường thẳng này, cho nửa đường thẳng phía ngoài [a, b] xuất phát từ a và cho nửa đường thẳng phía ngoài [a, b] xuất phát từ b.

Đa tạp tuyến tính trong R^n là tập $M \subset R^n$ sao cho M chứa cả đường thẳng đi qua hai điểm bất kì của nó. Chứng minh rằng M là đa tạp tuyến tính khác trống khi và chỉ khi $M = M_0 + \{m\}$, trong đó M_0 là một không gian con của R^n và $m \in R^n$ là một điểm của M. Khi đó M_0 sẽ gọi là *không gian con song song* (parallel subspace) với M. Chứng minh rằng cho trước đa tạp tuyến tính M thì không gian con M_0 song song với M là duy nhất, nhưng với $m^1, m^2 \in M$ bất kì thì

$$M_0 + \{m^1\} = M_0 + \{m^2\} = M.$$

4.3. *Số chiều* hoặc *thứ nguyên* (dimension) của đa tạp tuyến tính M, kí hiệu là dim M, là số chiều của không gian con song song của nó.

$$\{x = \lambda_0 a^0 + \lambda_1 a^1 + \ldots + \lambda_r a^r : \lambda_i \in R, \sum_{i=1}^{r} \lambda_i = 1\}$$

là đa tạp tuyến tính r chiều duy nhất đi qua r + 1 điểm $a^0, a^1, \ldots, a^r$ độc lập affine (xem Mục 4.6). (Đa tạp tuyến tính này gọi là *bao affine* (affine hull) của các điểm $a^0, a^1, \ldots, a^r$, kí hiệu là aff $\{a^0, a^1, \ldots, a^r\}$.

4.4. Cho $C \subset R^n$. *Bao affine của tập C* là bao affine của mọi hữu hạn các điểm của C, kí hiệu là aff(C). *Không gian con căng bởi C* (subspace spanned by C), hoặc *sinh bởi C* (subspace generated by C) hoặc *bao tuyến tính của C* (linear hull of C) là tổ hợp tuyến tính của mọi họ hữu hạn điểm của C, kí hiệu là span(C) hoặc lin(C). Chứng minh rằng

$$\text{span}(C) = \text{aff}(C \cup \{0\}). \tag{1}$$

Hãy biểu diễn hình học aff$\{a^0, \ldots, a^r\}$, aff(C), span(C) và (1). Chứng minh rằng aff(C) là đa tạp tuyến tính nhỏ nhất chứa C.

4.5. Chứng minh rằng tập $M \subset R^n$, khác trống, là đa tạp tuyến tính r chiều khi và chỉ khi nó có dạng

$$M = \{x \in R^n : Ax = b\},$$

ở đây A là ma trận cấp m × n nào đó, $b \in R^m$ có tính chất là rank A = rank (A, b) = n − r ≤ m.

4.6. *Siêu phẳng* (hyperplane) trong R^n là một đa tạp tuyến tính n − 1 chiều. Chứng minh rằng tập $P \subset R^n$ là siêu phẳng khi và chỉ khi nó có dạng

$$P = \{x \in R^n : a^T x = \alpha\},$$

ở đây $a \in R^n$, khác vectơ 0, và $\alpha \in R$.

4.7. Tập $C \subset R^n$ được gọi là *hình nón* (cone) nếu

$$x \in C, \lambda > 0 \Rightarrow \lambda x \in C. \tag{2}$$

(a) Nếu C là nón thì $x^0 + C$ được gọi là nón có mũi tại x^0. Hãy định nghĩa nón có mũi tại x^0 tương tự như (2) và biểu diễn hình học.

(b) Chứng minh rằng tập $C \subset R^n$ là nón lồi khi và chỉ khi

$$x, y \in C, \lambda > 0 \Rightarrow \lambda x \in C, x + y \in C$$

(c) Phát biểu và chứng minh khẳng định tương tự (b) cho nón lồi mũi tại x^0.

4.8. Tập con F của tập lồi C gọi là *diện* (face) của C nếu F lồi và mọi đoạn thẳng của C chứa một điểm $x \in F$ làm điểm trong của đoạn đều nằm trọn trong F. Ta gọi *số chiều của tập lồi C*, kí hiệu dim C, là số chiều của aff(C). Chứng minh rằng x là đỉnh (tức là điểm cực biên) của C khi và chỉ khi nó là diện của C và chỉ gồm một điểm.

Giao của mọi diện của C chứa x gọi là *diện nhỏ nhất* chứa x, kí hiệu F_x. Chứng minh rằng F_x là tập lồi gồm x và mọi $y \in C$ sao cho có đoạn $[y, z] \subset C$ nhận x là điểm trong của đoạn.

4.9. Chứng minh rằng số chiều (thứ nguyên) của tập lồi đa diện bằng n – r, ở đây r là hạng của ma trận ở các ràng buộc đẳng thức.

4.10. Giả sử tập lồi đa diện $P \subset R^n$ có dạng

$$\{x \in R^n: a_i^T x = b_i, i = 1,\ldots,m_1;\ a_i^T x \geq b_i, i = m_1 + 1,\ldots, m\}. \quad (3)$$

Chứng minh rằng tập $F \subset P$ là diện của P khi và chỉ khi nó có dạng

$$F = \{x: a_i^T x = b_i, i \in I;\ a_i^T x \geq b_i, i \in \{1,\ldots, m\} \backslash I\},$$

ở đây $I \supset \{1,\ldots, m_1\}$.

Từ đó suy ra:

(i) $x \in P$ là đỉnh khi và chỉ khi x thỏa chặt n ràng buộc độc lập tuyến tính;

(ii) một đoạn thẳng (hữu hạn hoặc vô hạn) là một cạnh của P khi và chỉ khi nó nằm trong P và thỏa chặt n – 1 ràng buộc độc lập tuyến tính.

4.11. Chứng minh rằng với mọi tập lồi đóng khác trống $C \subset R^n$ luôn tồn tại nón lồi đóng 0^+C sao cho $x + 0^+C \subset C$ $\forall x \in C$. 0^+C sẽ gọi là *nón lùi xa* (recession cone) của C, hoặc *nón tiệm cận* (asymptotic cone) của C, hoặc cũng gọi là *nón các hướng vô hạn* của C. Chứng minh rằng

(a) C giới nội khi và chỉ khi $0^+C = \{0\}$;

(b) $0^+C = \{d \in R^n: x + \lambda d \subset C\ \forall x \in C,\ \forall \lambda \geq 0\}$ (mỗi $d \in 0^+C$ là một hướng vô hạn của C);

(c) Nếu P là tập lồi đa diện (3) thì

$$0^+P = \{x \in R^n: a_i^Tx = 0,\ i = 1, \dots, m_1;\ a_i^Tx \geq 0,\ i = m_1 + 1, \dots, m\}.$$

4.12. Giả sử quy hoạch tuyến tính ở dạng chính tắc có tính chất là mọi nghiệm chấp nhận được đều thỏa $x_i \leq U$, $i = 1, \dots, n$, với một số U dương nào đó. Chứng minh rằng có thể đưa quy hoạch về dạng tương đương mà có ràng buộc $\sum_{i=1}^{n} x_i = 1$.

4.13. Xét phương pháp đơn hình theo hình học cột (Mục 4.6). Chứng minh rằng m + 1 điểm cơ sở $\begin{pmatrix} A_i \\ c_i \end{pmatrix}$ là độc lập affine.

4.14. Xét phương pháp đơn hình theo hình học cột. Chứng minh rằng khoảng cách thẳng đứng từ mặt phẳng đối ngẫu đến một điểm $\begin{pmatrix} A_i \\ c_i \end{pmatrix}$ bằng cước phí thu gọn (tức là hệ số mục tiêu) của biến x_i.

CHƯƠNG 5

HIỆU QUẢ CỦA PHƯƠNG PHÁP ĐƠN HÌNH. ĐỘ PHỨC TẠP TÍNH TOÁN

Ta đã biết là bằng phương pháp đơn hình có thể giải được sau hữu hạn bước mọi quy hoạch tuyến tính có nghiệm chấp nhận được (nếu dùng phép xoay chống xoay vòng). Bây giờ ta xét đến tốc độ hội tụ của phương pháp đơn hình.

5.1 SỐ BƯỚC LẶP ĐIỂN HÌNH

[Dantzig 1963] đã kết luận rằng thực tế sử dụng phương pháp đơn hình với các quy hoạch tuyến tính dạng chuẩn với số ràng buộc $m < 50$ và số biến n sao cho $m + n < 200$ cho thấy số bước lặp để đến nghiệm tối ưu thường không quá $3\frac{m}{2}$ và rất hiếm khi vượt quá $3m$. Thống kê gần đây trên các bài toán cỡ lớn cũng phù hợp: số bước lặp tăng tỉ lệ với m và tăng rất chậm theo n (khi m cố định, thì cỡ tăng tỉ lệ với $\log_e n$).

Chẳng hạn người ta đã giải thử trên 100 bài toán dạng chuẩn với $c_j = 1$ với mọi j, $b_i = 10\,000$ với mọi i và a_{ij} được lấy ngẫu nhiên trong tập các số nguyên từ 1 đến 1000. Quy tắc xoay được dùng là quy tắc Dantzig, cũng gọi là *quy tắc hệ số lớn nhất* (largest coefficient rule), tức là biến vào là biến có hệ số ở hàm mục tiêu âm và modul lớn nhất. Số bước lặp trung bình được cho trong bảng sau đây

m \ n	10	20	30	40	50
10	9,40	14,2	17,4	19,4	20,2
20		25,2	30,7	38,0	41,5
30			44,4	52,7	62,9
40				67,6	78,7
50					95,2

Bảng 5.1. Số bước lặp trung bình của quy tắc hệ số lớn nhất

5.2 TRƯỜNG HỢP XẤU NHẤT

Từ số liệu thống kê khả quan trong thực tế, người ta muốn chứng minh được một cách lí thuyết rằng mọi quy hoạch tuyến tính dạng chuẩn đều giải được bằng phương pháp đơn hình với, chẳng hạn, không quá 10 mn bước lặp.Tiếc là chưa ai làm được.[Klee - Minty 1972] đã tìm được phản thí dụ chứng tỏ rằng, về lí thuyết, phương pháp đơn hình tồi hơn người ta tưởng.

THÍ DỤ KLEE - MINTY.

$$\max \sum_{j=1}^{n} 10^{n-j} x_j,$$

$$2 \sum_{j=1}^{i-1} 10^{i-j} x_j + x_i \le 100^{i-1}, \quad i = 1, 2, ..., n,$$

$$x_j \ge 0, \qquad j = 1, 2, ..., n.$$

Với bài toán có m = n này, hai ông đã chỉ ra rằng nếu áp dụng quy tắc Dantzig thì sau đúng $2^n - 1$ bước lặp mới đi đến nghiệm tối ưu. Đây là số rất lớn (vì tăng theo hàm mũ!). Chẳng hạn với n = 50 thì máy tính với tốc độ tính 100 bước lặp một giây phải chạy 300.000 năm mới xong.

Để dễ hình dung, ta giải thử với n = 3:

$$\max\ 100x_1 + 10x_2 + x_3,$$
$$\begin{aligned} x_1 \qquad\qquad &\le 1, \\ 20x_1 + \quad x_2 \qquad &\le 100, \\ 200x_1 + 20x_2 + x_3 &\le 10\,000, \\ x_1, x_2, x_3 &\ge 0. \end{aligned}$$

Từ vựng đầu tiên sẽ là

$$\begin{aligned} z &= \qquad 100x_1 + 10x_2 + x_3 \\ \hline x_4 &= 1 - x_1 \\ x_5 &= 100 - 20x_1 - x_2 \\ x_6 &= 10000 - 200x_1 - 20x_2 - x_3. \end{aligned}$$

Áp dụng quy tắc Dantzig ta phải qua $7 = 2^3 - 1$ bước lặp như sau để đến nghiệm tối ưu.

Bước 1

$$\begin{aligned} z &= 100 - 100x_4 + 10x_2 + x_3 \\ \hline x_1 &= 1 - x_4 \\ x_5 &= 80 + 20x_4 - x_2 \\ x_6 &= 9800 + 200x_4 - 20x_2 - x_3. \end{aligned}$$

Bước 2

$$\begin{aligned} z &= 900 + 100x_4 - 10x_5 + x_3 \\ \hline x_1 &= 1 - x_4 \\ x_2 &= 80 + 20x_4 - x_5 \\ x_6 &= 8200 - 200x_4 + 20x_5 - x_3. \end{aligned}$$

Bước 3

$$\begin{array}{l}
\underline{z = 1000 - 100x_1 - 10x_5 + x_3} \\
x_4 = 1 - x_1 \\
x_2 = 100 - 20x_1 - x_5 \\
x_6 = 8000 + 200x_1 + 20x_5 - x_3.
\end{array}$$

Bước 4

$$\begin{array}{l}
\underline{z = 9000 + 100x_1 + 10x_5 - x_6} \\
x_4 = 1 - x_1 \\
x_2 = 100 - 20x_1 - x_5 \\
x_3 = 8000 + 200x_1 + 20x_5 - x_6.
\end{array}$$

Bước 5

$$\begin{array}{l}
\underline{z = 9100 - 100x_4 + 10x_5 - x_6} \\
x_1 = 1 - x_4 \\
x_2 = 80 + 20x_4 - x_5 \\
x_3 = 8200 - 200x_4 + 20x_5 - x_6.
\end{array}$$

Bước 6

$$\begin{array}{l}
\underline{z = 9900 + 100x_4 - 10x_2 - x_6} \\
x_1 = 1 - x_4 \\
x_5 = 80 + 20x_4 - x_2 \\
x_3 = 9800 + 200x_4 - 20x_2 - x_6.
\end{array}$$

Bước 7

$$\begin{array}{l}
\underline{z = 10000 - 100x_1 - 10x_2 - x_6} \\
x_4 = 1 - x_1 \\
x_5 = 100 - 20x_1 - x_2 \\
x_3 = 10000 - 200x_1 - 20x_2 - x_6.
\end{array}$$

Từ vựng đạt được sau bước 7 đã là tối ưu vì các hệ số ở hàng mục tiêu đều âm (đây là bài toán tìm max!). Nhận xét kĩ ta thấy nếu ở bước 1 ta chọn x_3 là biến vào thay cho x_1 thì sau một bước ta đến được ngay từ vựng tối ưu ở bước 7. Vì vậy thí dụ Klee - Minty có thể coi như một hiện tượng "bệnh lí" tức là xấu đặc biệt. Từ đây một câu hỏi được đặt ra tự nhiên là liệu có một quy tắc xoay khác để phương pháp đơn hình luôn chỉ cần một số nhỏ bước lặp. Thực ra quy tắc hệ số lớn nhất là khá tự nhiên từ cái nhìn ban đầu như đã thấy ở Mục 2.1, nhưng hãy xét kĩ hơn như sau. Giả sử ta thay đổi đơn vị đo của các biến ở thí dụ Klee - Minty theo quan hệ sau:

$$\bar{x}_1 = x_1;\ \bar{x}_2 = 0{,}01x_2;\ \bar{x}_3 = 0{,}0001x_3,$$

thì bài toán trở thành

$$\begin{array}{rl} \max 100\,\bar{x}_1 + 1000\bar{x}_2 + 10000\bar{x}_3, & \\ \bar{x}_1 & \le 1, \\ 20\bar{x}_1 + \ 100\bar{x}_2 & \le 100, \\ 200\bar{x}_1 + 2000\bar{x}_2 + 10000\bar{x}_3 & \le 10000, \\ \bar{x}_1, \bar{x}_2, \bar{x}_3 & \ge 0. \end{array}$$

Khi đó, cũng theo quy tắc hệ số lớn nhất bây giờ biến vào là x_3 và thuật toán kết thúc sau chỉ một bước lặp. Vậy là quy tắc Dantzig chưa tốt ở chỗ nó phụ thuộc vào đơn vị đo của biến chứ không chỉ phụ thuộc vào cấu trúc bản chất của bài toán. Ngay sau đây ta sẽ thấy một quy tắc tốt hơn.

5.3 QUY TẮC CẢI THIỆN NHIỀU NHẤT

Ta phải nhằm vào các quy tắc xoay độc lập với đơn vị đo các biến, tức là độc lập với biến đổi các biến bằng phép nhân tỷ lệ (khác nhau). Tất nhiên là phải chú ý đến mục đích chính là cải thiện hàm mục tiêu. Do đó hãy chọn biến vào làm cải thiện mục tiêu nhiều nhất.

Đây là *quy tắc cải thiện nhiều nhất* (largest improvement rule). Để thực hiện quy tắc này ở mỗi bước ta phải chọn biến vào, biến ra và tính ngay giá trị mục tiêu mới (tức là phần tử đầu của hàng đầu) rồi so sánh các giá trị này ở các ứng viên khác nhau. Vì vậy ở mỗi bước phải tính toán nhiều hơn. Nhưng ta thấy nếu giải bài toán Klee-Minty theo quy tắc này thì chỉ cần một bước, vì đương nhiên x_3 sẽ cho cải thiện lớn nhất (nó đã cho luôn giá trị mục tiêu lớn nhất mà !).

Cùng với bảng 5.1, [Avis - Chvátal 1978] cho bảng kết quả sau đây ứng với cùng các bài toán đã xét.

m \ n	10	20	30	40	50
10	7,02	9,17	10,8	12,1	12,6
20		16,2	20,2	24,2	27,3
30			28,7	34,5	39,4
40				43,3	49,9
50					58,9

Bảng 5,2 Số bước lặp trung bình của quy tắc cải thiện nhiều nhất

Ta thấy số bước lặp tương ứng ở bảng 5.2 nhỏ hơn hẳn ở Bảng 5.1. Thế nhưng trong thực tế sử dụng quy tắc hệ số lớn nhất lại tốn ít thời gian tính toán hơn quy tắc cải thiện nhiều nhất. (Vì mỗi bước lặp của quy tắc cải thiện nhiều nhất đòi hỏi tính toán nhiều.) Do đó tiêu chuẩn về số bước lặp ít cũng chỉ là một tiêu chuẩn "thô" để đánh giá hiệu quả của thuật toán. Trước khi đi sâu hơn về hiệu quả của thuật toán ta hãy dừng lại minh họa hình học thí dụ Klee - Minty để thấy rõ hơn rằng áp dụng quy tắc hệ số lớn nhất ở đây cho ta "cách đi xấu nhất". Với n = 3 ta thấy miền chấp nhận được xấp xỉ khối hộp chữ nhật

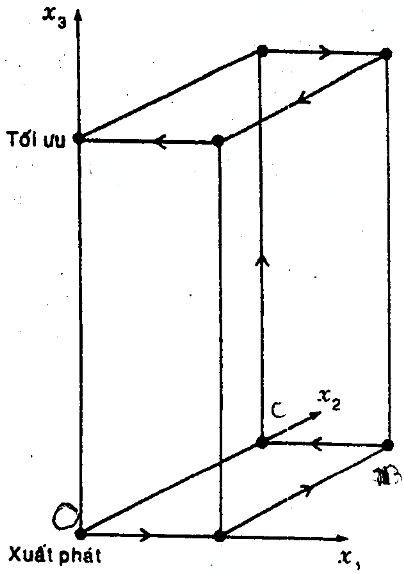
$0 \le x_1 \le 1,$

$0 \le x_2 \le 100,$

$0 \le x_3 \le 10000,$

(vì có thể coi $x_1 << x_2 << x_3$). Khối này có $2^3 = 8$ đỉnh. Ta xuất phát từ đỉnh (0, 0, 0) (ứng với từ vựng ban đầu) mà phải qua $7 = 2^3 - 1$ đỉnh, tức là phải qua tất cả các đỉnh để đến được đỉnh tối ưu. Mà đỉnh này lại kề ngay đỉnh xuất phát. Ta có thể thấy điều này từ H.5.1 hoặc từ nhận xét giải tích: đỉnh tối ưu có thể đi từ đỉnh xuất phát chỉ bởi một bước lặp (với x_3 là biến vào) nên phải là đỉnh kề.

Trường hợp n chiều, tức là n biến cũng hoàn toàn như vậy.



H. 5.1 Biểu diễn xấp xỉ thí dụ Klee -Minty với n=3

Một câu hỏi được đặt ra: nếu theo tiêu chuẩn về số bước lặp liệu quy tắc cải thiện nhiều nhất có nhiều hứa hẹn không? Nếu xét đến trường hợp xấu nhất thì [Jeroslow 1973] đã cho câu trả lời phủ định bởi một thí dụ kiểu Klee - Minty, tức là chỉ ra trường hợp số bước lặp theo quy tắc cải thiện nhiều nhất tăng theo hàm mũ với m và n.

5.4 ĐỘ PHỨC TẠP TÍNH TOÁN

Ở trên ta đã đánh giá hiệu quả của thuật toán đơn hình bằng số bước lặp tăng nhanh hay chậm theo m và n, được coi là đặc trưng cho cỡ của bài toán. Đồng thời ta cũng đã nói đến "trường hợp xấu nhất" và "đánh giá trung bình" đến hiệu quả của thuật toán tính bởi thời gian thực hiện tính toán trên máy tính. Ta cũng đã thấy đó cũng chỉ là những yếu tố đánh giá được phần nào hiệu quả tính toán.

Hiệu quả của thuật toán và *độ phức tạp tính toán* (computational complexity) là những vấn đề lớn của toán học và tin học, chứa nhiều quan niệm và định nghĩa đến nay còn đang tiếp tục được bàn luận.

5.4.1 Đánh giá hiệu quả

Có hai cách đánh giá hiệu quả: trường hợp xấu nhất và trường hợp trung bình. Phân tích trường hợp xấu nhất là xét tất cả các bài toán của cùng một "cỡ" và xem cần bao nhiêu nỗ lực tính toán để giải bài toán khó nhất trong đó. Phân tích trường hợp trung bình là xem nỗ lực trung bình cho tất cả các bài toán cùng cỡ.

Phân tích trường hợp xấu nhất dễ hơn vì ta chỉ cần cho một cận trên cho tất cả các nỗ lực tính toán rồi đưa ra một thí dụ đặc biệt đạt được cận trên đó. Trái lại, để phân tích trường hợp trung bình ta cần một mô hình thống kê, ngẫu nhiên, là một việc rất khó.

Tuy vậy, phân tích trường hợp xấu nhất thường chỉ có ý nghĩa lí thuyết, còn phân tích trường hợp trung bình mới có ý nghĩa đối với người giải các bài toán thực tế. Chính vì vậy, Klee - Minty và Jeroslow đã cho thấy ở các trường hợp xấu thì số bước lặp cần thiết của phương pháp đơn hình là tăng theo hàm mũ theo cỡ của bài toán, nhưng thực tế phương pháp đơn hình vẫn cho lời giải rất nhanh.

5.4.2 Cỡ của bài toán

Như trên đã thấy hai tham số m và n được dùng một cách tự nhiên để đặc trưng cho *cỡ* (size) của bài toán. Đây luôn là một cách đặc trưng phổ biến trong thực tế. Nhưng cũng nảy sinh nhiều điều cần cân nhắc. Một là, thông thường người ta cần một tham số duy nhất đặc trưng cho cỡ bài toán, số này cũng có thể lấy là m.n (vì có m.n hệ số của các ràng buộc). Đặc trưng này của "cỡ" cũng hay được dùng. Xét kĩ hơn trong thực tế, rất nhiều hoặc thậm chí hầu hết các hệ số bằng 0 và các thuật toán tốt cần tận dụng điều này và thực tế đã giảm được khối lượng tính toán nhiều. Vì thế "cỡ" có thể đặc trưng bởi số các dữ kiện khác 0 sẽ chính xác hơn.

Người ta lại có thể đi xa hơn và nhận xét rằng một dữ kiện đơn giản như 5 hay 12 thì đưa vào máy và xử lí nhanh hơn một dữ kiện nhiều chữ số, chẳng hạn 325137, 4831. Vì vậy, đa số các nhà tin học dùng số L các bít cần để lưu toàn bộ dữ kiện của bài toán vào máy tính đặc trưng cho cỡ của bài toán.

5.4.3 Đánh giá khối lượng tính toán

Đánh giá khối lượng tính toán, tức là toàn bộ công việc để giải một bài toán bằng thuật toán, bởi đại lượng nào?

Có thể câu trả lời hợp lí là: bởi thời gian máy tính chạy.

Nhưng lại có vấn đề là mỗi người dùng một máy tính khác nhau. Hơn nữa kĩ thuật máy tính thay đổi rất nhanh. Sau vài năm có thể thế hệ máy khác được sử dụng. Vì vậy thời gian máy chạy có vẻ là một đặc trưng hứa hẹn cho khối lượng tính toán, nhưng lại không được thực tế lắm. Người ta vẫn hay dùng số bước lặp để đánh giá khối lượng này vì nó không phục thuộc máy tính, cũng không phụ thuộc ngôn ngữ lập trình. Mặc dù rằng ở các lớp thuật toán khác nhau thì thời gian chạy máy của một bước lặp có thể rất khác nhau, nhưng trong cùng một lớp thì thời gian này có thể coi là tương tự.

Cũng có một cách đánh giá nữa là dùng số phép toán (operation) chứa trong thuật toán .(Ở đây có thể coi là phép toán số học, nhưng chỉ tính đến số phép nhân và chia, vì phép cộng được thực hiện trên máy quá nhanh so với nhân.) Tổng quát hơn, người ta dùng số chỉ dẫn (instruction) để thực hiện thuật toán thay cho số phép toán hoặc số bước lặp. Các đại lượng này đều độc lập với máy và ngôn ngữ.

5.4.4 Độ phức tạp của thuật toán

ĐỊNH NGHĨA 5.4.1. *Độ phức tạp* (complexity) của một thuật toán A là hàm $t_A(S)$ cho ta số các chỉ dẫn để thực hiện thuật toán cho trường hợp xấu nhất trong tất cả các bài toán, mà thuật toán A có thể dùng, có cỡ S.

Như vậy ở đây ta đã cố định lớp bài toán(chẳng hạn là lớp bài toán quy hoạch tuyến tính). Cỡ S được hiểu tổng quát hoặc là một trong các định nghĩa cỡ cụ thể ở trên.

Để so sánh độ phức tạp tính toán $t_A(S)$ của các thuật toán khác nhau,cho cùng một bài toán ta dùng khái niệm cấp của một hàm sau đây:

Cho hai hàm số f và g từ R vào R. Ta nói f có cấp số nhỏ hơn hoặc bằng g, và viết là $f = 0(g)$, nếu: $\exists x_0, \exists c > 0, \forall x > x_0, f(x) \leq cg(x)$.

Bảng sau đây cho thấy thời gian tính bằng máy PC 486 các thuật toán với một số độ phức tạp tính toán điển hình. Ta thấy là với cỡ không quá nhỏ thì thuật toán có độ phức tạp tính toán không phải là đa thức (ở nửa dưới của bảng) đòi hỏi thời gian quá lớn, không dùng được trong thực tế. Chẳng hạn trong quản lí kinh doanh, để quyết định chiến lược cạnh tranh lãnh đạo công ti thường cần có kết quả tính toán ngay trong một ngày (< 30 giờ).

cỡ S / $t_A(S)$	20	50	100	200	500	1000
10^3S	0,02s	0,05s	0,1s	0,2s	0,5s	1s
$10^3S\log_2S$	0,09s	0,3s	0,6s	1,5s	4,5s	10s
$100S^2$	0,04s	0,25s	1s	4s	25s	2m
$10S^3$	0,02s	1s	10s	1m	21m	27h
$S^{\log 2S}$	0,4s	1,1h	220d	12500y	5.10^{10}y	-
$2^{\frac{S}{3}}$	0,0001s	0,1s	2,7h	3.10^6y	-	-
2^S	1s	36y	-	-	-	-
3^S	58m	2.10^{11}y	-	-	-	-
S!	77100y	-	-	-	-	-

Bảng 5.3 Thời gian tính trên PC 486 của các thuật toán

s: giây, m: phút, h: giờ, d: ngày, y: năm, - có nghĩa là số lớn quá.

Một thuật toán A được gọi là *thuật toán vừa ý* (satisfactory algorithm) nếu có một đa thức p(S) sao cho $t_A = 0(p)$, tức là $\exists S_0, \exists c, \forall S > S_0, t_A(S) \leq c.p(S)$. Khái niệm được đưa ra bởi [Edmonds 1965] này rất được hưởng ứng và sử dụng rộng rãi trong toán ứng dụng và tin học.

Thuật toán vừa ý cũng gọi là *thuật toán đa thức* (polynominal algorithm).

5.4.5 Độ phức tạp của bài toán

Độ phức tạp của bài toán là vấn đề khác với độ phức tạp của thuật toán vì một bài toán có nhiều thuật toán giải.

Ở đây lí thuyết về độ phức tạp chỉ tập trung trên một loại bài toán là bài toán về tồn tại.

Bài toán tồn tại (existence problem) là bài toán tìm một phần tử nào đó có một tính chất cho trước trong một tập hữu hạn cho trước. Vì vậy lời giải của nó mang tính chất là "có" hay "không".

Các bài toán không phải ở dạng bài toán tồn tại muốn được xét độ phức tạp tính toán phải được đưa về dạng bài toán tồn tại. Chẳng hạn, một bài toán tối ưu có thể đưa về dãy P_k các bài toán tồn tại: có tồn tại nghiệm thỏa các ràng buộc của bài toán tối ưu và cho giá trị mục tiêu tương ứng là k? Ở đây phải cố đưa tập ràng buộc về dạng hữu hạn. Thí dụ đối với quy hoạch tuyến tính thì đây là tập các đỉnh của miền chấp nhận được.

Bài toán dễ (easy problem) là bài toán tồn tại mà có thuật toán đa thức để giải nó.

THÍ DỤ 5.4.1. Trong các bài toán nổi tiếng, các bài toán sau đây chẳng hạn là các bài toán dễ:

- Quy hoạch tuyến tính (với các thuật toán điểm trong);
- Bài toán đường ngắn nhất trong mạng (có $t_A(S) = 0(S^2)$, xem Mục 11.4.2);

- Bài toán dòng cực đại trong mạng (có $t_A(S) = 0(n.m^2)$, xem Mục 10.7.2).

Lớp tất cả các bài toán dễ thường gọi là *lớp P* (từ thuật ngữ polynomial).

Lý thuyết về độ phức tạp thường tập trung trên thực tế áp dụng. Giả sử ta đã dùng thuật toán nào đó hoặc bằng cách nào đó cho được nghiệm của bài toán tồn tại để giám đốc Công ti dùng để quyết định chiến lược cạnh tranh. Vì sự cạnh tranh là sống còn ông giám đốc khôn ngoan và thực tế thường đòi hỏi ta phải chỉ ra cách kiểm tra xem có đúng là nghiệm hay không. Vì vậy có định nghĩa sau đây.

Một bài toán tồn tại được gọi là thuộc *lớp NP* (từ thuật ngữ nondeterministic polynomial) nếu có một thuật toán đa thức để cho nghiệm của nó hoặc để kiểm tra xem một phương án đưa ra có đúng là nghiệm không.

Như vậy rõ ràng $P \subset NP$.

Một lớp bài toán khác là *lớp NP - khó* (NP - hard hoặc NP complete). Một bài toán là thuộc lớp NP - khó nếu loài người chưa biết thuật toán đa thức nào để giải nó.

THÍ DỤ 5.4.2. Các bài toán sau đây là NP - khó

- Bài toán người giao hàng (travelling salesman problem), xem Thí dụ 12.5.1;
- Bài toán quy hoạch tuyến tính nguyên (integer linear programming);
- Bài toán tô mầu đồ thị (minimum coloring).

Chú ý rằng ta chưa biết thuật toán đa thức để giải thì cũng có thể sẽ tìm được thuật toán như vậy. Vì thế đang có bài toán mở rất lớn là $P \supset NP$ - khó hay $P \not\supset NP$ - khó? Người ta phỏng đoán khả năng thứ hai.

BÀI TẬP CHƯƠNG 5

Xét quy hoạch tuyến tính, với $\varepsilon \in \left(0, \frac{1}{2}\right)$,

$$\max x_n,$$
$$\varepsilon \le x_1 \le 1,$$
$$\varepsilon x_{i-1} \le x_i \le 1 - \varepsilon x_{i-1},\ i = 2,\dots, n.$$

Chứng minh các khẳng định sau

(a) tập chấp nhận được có 2^n đỉnh;

(b) các đỉnh có thể được sắp thứ tự sao cho mỗi đỉnh là kề và có giá trị mục tiêu lớn hơn đỉnh trước nó;

(c) có một quy tắc xoay sao cho phương pháp đơn hình tương ứng cần $2^n - 1$ lần đổi cơ sở mới kết thúc.

CHƯƠNG 6

LÝ THUYẾT ĐỐI NGẪU

Lý thuyết đối ngẫu là một vấn đề trung tâm của lí thuyết tối ưu. Nó giúp phân tích sâu cấu trúc của bài toán tối ưu, cho phép giải thích hình học, đồng thời cũng là công cụ mạnh mẽ để nghiên cứu và xây dựng các thuật toán. Nói riêng, trong quy hoạch tuyến tính, ở chương này ta cũng thấy rõ điều đó.

6.1 BÀI TOÁN MỞ ĐẦU

Xét một quy hoạch tuyến tính, chẳng hạn ở dạng chính tắc

$$\begin{aligned} &\min c^T x, \\ &Ax = b, \\ &x \geq 0. \end{aligned} \tag{1}$$

Nếu ta biết một nghiệm chấp nhận được x^0 thì rõ ràng ta được một cận trên của mục tiêu tối ưu. Giả sử là có tồn tại nghiệm tối ưu x^* thì $c^T x^* \leq c^T x^0$. Tuy ta chưa tìm được x^*, nhưng nếu ta biết một cận dưới của mục tiêu tối ưu thì đã "khoanh vùng" được giá trị mục tiêu tối ưu. Ta hãy thử tìm một cận dưới.

Xử lí ràng buộc $Ax = b$ là khó nên ta hãy thử "nới lỏng" bằng cách coi như không có ràng buộc này nhưng phạt vào các phương án x không thỏa ràng buộc này (tức $Ax - b \neq 0$) một lượng $y^T(b - Ax)$, ở đây y là véc tơ (m chiều như b) giá phải trả cho sự vi phạm ràng buộc $b - Ax$. Thay cho (1) ta giải bài toán

$$\min c^T x + y^T(b - Ax),$$
$$x \geq 0. \qquad (2)$$

y được gọi là *nhân tử Lagrange* (Lagrange multiplier), $L(x,y) := c^T x + y^T(b - Ax)$ được gọi là *hàm Lagrange* (Lagrange function) hoặc *Lagrangean* của bài toán (1).

Đặt g(y) là giá trị tối ưu của bài toán nới lỏng (2) (phụ thuộc vectơ giá y). Vì bài toán nới lỏng bị ràng buộc ít hơn nên ta dự đoán g(y) là cận dưới cho giá trị mục tiêu tối ưu $c^T x^*$ của bài toán (1). Điều này được kiểm chứng dễ dàng:

$$g(y) = \min_{x \geq 0}(c^T x + y^T(b - Ax)) \leq c^T x^* + y^T(b - Ax^*) = c^T x^*,$$

vì x^* là chấp nhận được của (1) nên $b - Ax^* = 0$. Vậy ta đã có một cận dưới cho giá trị mục tiêu tối ưu của (1) là g(y) với vec tơ giá y bất kì. Một cách tự nhiên ta quan tâm bài toán tìm cận dưới sát nhất:

$$\max_{y \in R^m} g(y), \qquad (3)$$

Bài toán này gọi là *bài toán đối ngẫu* (dual problem) của quy hoạch tuyến tính (1). Sau đây, ta sẽ thấy kết quả chính của lí thuyết đối ngẫu của quy hoạch tuyến tính khẳng định giá trị mục tiêu tối ưu của bài toán gốc (primal problem), tức là bài toán (1), bằng giá trị mục tiêu tối ưu của bài toán đối ngẫu (3) của nó. Nói một cách khác, nếu vec tơ giá y được chọn thích hợp là nghiệm tối ưu của bài toán đối ngẫu thì việc vi phạm ràng buộc $b = Ax$ không làm ảnh hưởng đến giá trị mục tiêu tối ưu (vì giá trị mục tiêu tối ưu của bài toán nới lỏng (2) bằng $c^T x^*$).

Bây giờ ta hãy biến đổi (3) để được dạng tốt hơn. Ta có

$$g(y) = \min_{x \geq 0} (c^T x + y^T(b - Ax)) = y^T b + \min_{x \geq 0} (c^T - y^T A)x.$$

Mặt khác

$$\min_{x \geq 0} (c^T - y^T A)x = \begin{cases} 0, & \text{nếu } c^T - y^T A \geq 0^T, \\ -\infty & \text{nếu khác.} \end{cases}$$

Thay vào (3) ta được dạng tương đương của nó

$$\max y^T b, \quad y^T A \leq c^T \tag{4}$$

Vậy bài toán đối ngẫu của quy hoạch tuyến tính (1) là quy hoạch tuyến tính (4) được lập dễ dàng từ các dữ kiện của (1).

Ta lập bài toán đối ngẫu cho quy hoạch tuyến tính dạng chuẩn một cách tương tự, bằng cách thêm biến bù

$$Ax \leq b \quad \Leftrightarrow \quad Ax + s = b,\ s \geq 0.$$

$$\Leftrightarrow \quad [A \ \ I]\begin{bmatrix} x \\ s \end{bmatrix} = b,\ s \geq 0.$$

Vì với biến bù, mục tiêu trở thành $\min c^T x + 0^T s$, nên theo (4), áp dụng cho trường hợp có biến bù này ta được bài toán đối ngẫu của quy hoạch tuyến tính dạng chuẩn là

$$\max y^T b, \quad y^T [A \ I] \leq [c^T \ 0^T];$$

tức là

$$\max y^T b, \quad y^T A \leq c^T, \quad y \leq 0. \tag{5}$$

Bây giờ ta xét trường hợp bài toán gốc không ở dạng chính tắc hoặc dạng chuẩn. Khi đó để lập bài toán đối ngẫu, tức là bài toán tìm cận dưới sát nhất của giá trị mục tiêu tối ưu ta cũng lí luận tương tự như trên. Với các ràng buộc đẳng thức, bất đẳng thức và ràng buộc về dấu ta làm đúng như trên. Nếu vec tơ x ở bài toán gốc là tự do (về dấu) thì cũng lí luận như với (4) ta có

$$\min_{x \in R^n}(c^T - y^T A)x = \begin{cases} 0, & \text{nếu } c^T - y^T A \geq 0^T, \\ -\infty & \text{nếu khác.} \end{cases}$$

Vậy ở bài toán đối ngẫu phải có ràng buộc

$$y^T A = c^T.$$

6.2 ĐỊNH NGHĨA BÀI TOÁN ĐỐI NGẪU

Giả sử ma trận A có hàng là a_i^T và các cột là A_j. Bởi lí luận ở Mục 6.1, ta dẫn đến định nghĩa tổng quát của bài toán đối ngẫu như sau. Giả sử bài toán gốc có cấu trúc ở bên trái trong bảng dưới dây, Khi đó bài toán đối ngẫu được định nghĩa với cấu trúc tương ứng ở bên phải

$\min c^T x$,	$\max y^T b$,
với ràng buộc	với ràng buộc
$a_i^T x = b_i,\ i \in M_1$,	y_i tự do, $i \in M_1$,
$a_i^T x \leq b_i,\ i \in M_2$,	$y_i \leq 0,\ i \in M_2$,
$a_i^T x \geq b_i,\ i \in M_3$,	$y_i \geq 0,\ i \in M_3$,
$x_j \geq 0,\ j \in N_1$,	$y^T A_j \leq c_j^T,\ j \in N_1$,
$x_j \leq 0,\ j \in N_2$,	$y^T A_j \geq c_j^T,\ j \in N_2$,
x_j tự do, $j \in N_3$,	$y^T A_j = c_j^T,\ j \in N_3$.

NHẬN XÉT. Mỗi ràng buộc ở bài toán gốc, không kể ràng buộc dấu, ứng với một biến của bài toán đối ngẫu. Mỗi biến của bài toán gốc ứng với một ràng buộc ở bài toán đối ngẫu. Đồng thời các chiều bất đẳng thức có quan hệ trực tiếp với nhau và cho bằng bảng sau đây.

GỐC	min	max	ĐỐI NGẪU
ràng buộc	$= b_i$	tự do	biến
	$\leq b_i$	≤ 0	
	$\geq b_i$	≥ 0	
biến	≥ 0	$\leq c_j$	ràng buộc
	≤ 0	$\geq c_j$	
	tự do	$= c_j$	

Bảng 6.1 Quan hệ giữa biến và ràng buộc của bài toán gốc và bài toán đối ngẫu

Các biến y_i của bài toán đối ngẫu được gọi là *biến đối ngẫu* (dual variable). Ở trên ta đã thấy là biến đối ngẫu cũng chính là nhân tử Lagrange.

Mỗi quy hoạch tuyến tính có nhiều dạng tương đương. Từ mỗi dạng đó, theo định nghĩa trên ta sẽ lập được một bài toán đối ngẫu. Tất nhiên các bài toán đối ngẫu nhận được cũng sẽ tương đương nhau.

Một câu hỏi được đặt ra: bài toán đối ngẫu của bài toán đối ngẫu có khác bài toán gốc không ?

THÍ DỤ 6.2.1. Xét quy hoạch tuyến tính ở bên trái dưới đây ta sẽ lập được bài toán đối ngẫu của nó như ở bên phải.

$$\begin{array}{ll}
\min & x_1 + x_2 + 3x_3, \\
& -x_1 + 3x_2 = 5, \\
& 2x_1 - x_2 + 3x_3 \geq 6, \\
& x_3 \leq 4, \\
& x_1 \geq 0, \\
& x_2 \leq 0, \\
& x_3 \text{ tự do.}
\end{array}
\qquad
\begin{array}{ll}
\max & 5y_1 + 6y_2 + 4y_3, \\
& y_1 \text{ tự do,} \\
& y_2 \geq 0, \\
& y_3 \leq 0, \\
& -y_1 + 2y_2 \leq 1, \\
& 3y_1 - y_2 \geq 1, \\
& 3y_2 + y_3 = 3.
\end{array}$$

Bây giờ ta tìm đối ngẫu của bài toán đối ngẫu trên đây. Ta viết lại nó ở dạng bài toán min và nhân cả ba ràng buộc cuối với -1. Dạng tương đương nhận được ta viết vào bên trái dưới đây và ta lập đối ngẫu của nó ở bên phải.

$$\begin{array}{ll}
\min -5y_1 - 6y_2 - 4y_3, & \max -x_1 - x_2 - 3x_3, \\
y_1 \text{ tự do}, & x_1 - 3x_2 = -5, \\
y_2 \geq 0, & -2x_1 + x_2 - 3x_3 \leq -6, \\
y_3 \leq 0, & -x_3 \geq -4, \\
y_1 - 2y_2 \geq -1, & x_1 \geq 0, \\
-3y_1 + y_2 \leq -1, & x_2 \leq 0, \\
-3y_2 - y_3 = -3. & x_3 \text{ tự do}.
\end{array}$$

Ta thấy bài toán đối ngẫu ở bên phải trên đây của bài toán đối ngẫu trùng với bài toán gốc.

Bài toán gốc ta xét ở thí dụ trên đã có đủ các kiểu ràng buộc. Vì vậy làm đúng như ở thí dụ trên đối với bài toán tổng quát (chứ không phải bằng số cụ thể) ta chứng minh được định lí sau đây

ĐỊNH LÍ 6.2.1. Nếu ta biến đổi bài toán đối ngẫu về dạng tương đương là tìm cực tiểu (min) và lập đối ngẫu của nó ta nhận được bài toán gốc.

6.3 CÁC ĐỊNH LÍ VỀ ĐỐI NGẪU

6.3.1 Đối ngẫu yếu

Ở Mục 6.1 ta đã thấy đối với bài toán chính tắc thì giá trị mục tiêu g(y) của bài toán đối ngẫu luôn là cận dưới của giá trị mục tiêu tối ưu của bài toán gốc. Định lí sau cho thấy tính chất này đúng với bất kì cặp bài toán tuyến tính đối ngẫu dạng bất kì.

ĐỊNH LÍ 6.3.1 (đối ngẫu yếu). Nếu x là nghiệm chấp nhận được của quy hoạch tuyến tính gốc và y là nghiệm chấp nhận được của quy *hoạch đối ngẫu thì*

$$y^T b \le c^T x.$$

CHỨNG MINH. Ta đặt

$$u_i = y_i(a_i^T x - b_i),$$
$$v_j = (c_j - y^T A_j)x_j.$$

Theo định nghĩa bài toán đối ngẫu (Bảng 6.1) thì y_i và $a_i^T x - b_i$ cùng dấu, còn $c_j - y^T A_j$ và x_j cùng dấu. Do đó $u_i \ge 0 \ \forall i$ và $v_j \ge 0 \ \forall j$.

Nhận xét rằng

$$\sum_i u_i = y^T A x - y^T b,$$
$$\sum_j v_j = c^T x - y^T A x.$$

Do đó

$$0 \le \sum_i u_i + \sum_j v_j = c^T x - y^T b. \ \square$$

HỆ QUẢ 6.3.2.

(i) Nếu hàm mục tiêu của quy hoạch tuyến tính gốc không giới nội dưới, thì bài toán đối ngẫu không có nghiệm chấp nhận được.

(ii) Nếu bài toán đối ngẫu không giới nội trên (tức là hàm mục tiêu của nó không giới nội trên), thì bài toán gốc không chấp nhận được (tức là không có nghiệm chấp nhận được).

CHỨNG MINH. Do sự tương tự ta chỉ chứng minh (i). Giả sử bài toán gốc không giới nội dưới, tức là tồn tại dãy x^k các nghiệm chấp nhận được để $c^T x^k \to -\infty$ khi $k \to \infty$. Giả sử phản chứng là bài toán đối ngẫu có nghiệm chấp nhận được y^0. Khi đó, do đối ngẫu yếu,

$$(y^0)^T b \le c^T x^k \ \forall k.$$

Cho $k \to \infty$ ta được điều vô lí $(y^0)^T b \le -\infty$. □

HỆ QUẢ 6.3.3. Giả sử $\bar{x}$ và $\bar{y}$ là nghiệm chấp nhận được tương ứng của hai bài toán đối ngẫu nhau. Nếu $\bar{y}^T b = c^T \bar{x}$ thì $\bar{x}$ và $\bar{y}$ phải là nghiệm tối ưu tương ứng của hai bài toán.

CHỨNG MINH. Xét nghiệm chấp nhận được x bất kì của bài toán gốc. Theo đối ngẫu yếu ta có $\bar{y}^T b \le c^T x$, và do đó $c^T \bar{x} \le c^T x$, $\forall x$ chấp nhận được. Vậy $\bar{x}$ là nghiệm tối ưu. Tương tự ta sẽ thấy y là nghiệm tối ưu của bài toán đối ngẫu.□

6.3.2 Đối ngẫu mạnh

Bây giờ ta chứng minh kết quả chính nhất của lí thuyết đối ngẫu đã có nhắc đến ở Mục 6.1 là giá trị mục tiêu tối ưu của hai bài toán quy hoạch tuyến tính đối ngẫu nhau là bằng nhau.

ĐỊNH LÍ 6.3.4 (đối ngẫu mạnh). Nếu quy hoạch tuyến tính gốc có nghiệm tối ưu x^* thì quy hoạch đối ngẫu cũng có nghiệm tối ưu y^* và giá trị mục tiêu tối ưu bằng nhau:

$$c^T x^* = (y^*)^T b.$$

CHỨNG MINH. Quy hoạch tuyến tính có nhiều dạng tương đương, nhưng ta đã biết là các bài toán đối ngẫu của các dạng tương đương đó cũng là các dạng tương đương. Mà các bài toán tương đương thì có cùng giá trị mục tiêu tối ưu. Vì vậy ta chỉ cần chứng minh định lí cho một dạng cụ thể, chẳng hạn là dạng chuẩn:

$$\begin{aligned} &\min c^T x, \\ &Ax \le b, \\ &x \ge 0. \end{aligned}$$

Ta sẽ áp dụng phương pháp đơn hình lên quy hoạch này. Vì có nghiệm tối ưu nên sau hữu hạn bước ta sẽ được từ vựng tối ưu với hàm mục tiêu là

$$z = z^* + \sum_{j=1}^{n} c_j^* x_j + \sum_{i=1}^{m} d_i^* w_i, \tag{6}$$

ở đây, $c_j^* = 0$ và $d_i^* = 0$ với các biến x_j và w_i là biến cơ sở ở từ vựng tối ưu này. Vì z^* là giá trị mục tiêu tối ưu nên (nó phải tương ứng với nghiệm tối ưu)

$$z^* = \sum_{j=1}^{n} c_j x_j^* . \tag{7}$$

Đặt

$$y_i^* = - d_i^*, i = 1, ..., m.$$

Ta sẽ chứng minh $y^* = (y_1^*, ..., y_m^*)$ là nghiệm chấp nhận được của bài toán đối ngẫu với $(y^*)^T b = c^T x^*$. Vì hàm mục tiêu ban đầu và (6) là tương đương, ta có

$$\begin{aligned}\sum_{j=1}^{n} c_j x_j &= z^* + \sum_{j=1}^{n} c_j^* x_j + \sum_{i=1}^{m} d_i^* w_i \\ &= z^* + \sum_{j=1}^{n} c_j^* x_j + \sum_{i=1}^{m} (-y_i^*)(b_i - \sum_{j=1}^{n} a_{ij} x_j) \\ &= z^* - \sum_{i=1}^{m} b_i y_i^* + \sum_{j=1}^{n} (c_j^* + \sum_{i=1}^{m} y_i^* a_{ij}) x_j .\end{aligned}$$

Hai vế là biểu diễn của cùng một hàm tuyến tính nên các hệ số tương ứng phải đồng nhất. Do đó

$$z^* = \sum_{i=1}^{m} b_i y_i^*, \tag{8}$$

$$c_j = c_j^* + \sum_{i=1}^{m} y_i^* a_{ij}. \tag{9}$$

Từ (7) và (8) ta có $(y^*)^T b = c^T x^*$. Ở từ vựng tối ưu các hệ số của hàm mục tiêu phải không âm, tức $c_j^* \geq 0$, $j = 1, \ldots, n$, và $d_i^* \geq 0$, $i = 1, \ldots, m$. Vì vậy, theo (9)

$$\sum_{i=1}^{m} y_i^* a_{ij} \leq c_j, \; j = 1, \ldots, n,$$

$$y_i^* = -d_i^* \leq 0, \; i = 1, \ldots, m,$$

tức là y^* là nghiệm chấp nhận được của bài toán đối ngẫu. Do $(y^*)^T b = c^T x^*$, y^* phải là nghiệm tối ưu theo Hệ quả 6.3.3.□

NHẬN XÉT 6.3.1

(a) Sau định lí đối ngẫu yếu ta mới chỉ biết $(y^*)^T b \leq (c)^T x^*$ (nếu hai bài toán có nghiệm tối ưu). Khi xét các cặp bài toán đối ngẫu nhau cho quy hoạch phi tuyến tổng quát người ta cũng đạt được sự so sánh giá trị mục tiêu tối ưu ở dạng đối ngẫu yếu như vậy. Nếu dấu bất đẳng thức chặt xảy ra ta nói có *lỗ hổng đối ngẫu* (duality gap). Đối với quy hoạch phi tuyến muốn lỗ hổng này bằng 0;bài toán gốc phải là bài toán lồi, tức là các tập và hàm trong dữ liệu phải thỏa các giả thiết nhất định về lồi. Định lí 6.3.4 về đối ngẫu mạnh nói rằng đối với quy hoạch tuyến tính chấp nhận được thì lỗ hổng đối ngẫu luôn bằng 0. Điều này dễ hiểu vì quy hoạch tuyến tính là trường hợp riêng của quy hoạch lồi.

(b) Hệ quả 6.3.2 có thể "ghép vào" Định lí 6.3.4 như sau. Quy hoạch gốc không giới nội được quy ước (khi nói đến đối ngẫu) như là mục tiêu tối ưu bằng $-\infty$. Khi đó hệ quả nói rằng quy hoạch đối ngẫu không chấp nhận được. Điều này có thể quy ước coi là giá trị mục tiêu của bài toán đối ngẫu (bài toán max) luôn bằng $-\infty$, Như vậy ta cũng coi như lỗ hổng đối ngẫu bằng 0, tức là ở dạng đối ngẫu mạnh. Tương tự, trường hợp bài toán đối ngẫu không giới nội có thể coi như giá trị mục tiêu tối ưu của bài toán bằng nhau ở $+\infty$.

(c) Theo quy ước mở rộng cách nói như ở (b) liệu có phải quy hoạch tuyến tính(không cần chấp nhận được) luôn có lỗ hổng đối ngẫu bằng 0 hay không? Thí dụ sau đây cho câu trả lời phủ định.

THÍ DỤ 6.3.1. Cặp quy hoạch đối ngẫu sau đây cùng không chấp nhận được, tức là không có nghiệm chấp nhận được,

$$\begin{array}{ll} \min -2x_1 + x_2, & \max y_1 - 2y_2, \\ x_1 - x_2 \le 1, & y_1 \le 0, \\ -x_1 + x_2 \le -2, & y_2 \le 0, \\ x_1 \ge 0, & y_1 - y_2 \le -2, \\ x_2 \ge 0. & -y_1 + y_2 \le 1. \end{array}$$

Trường hợp này có thể quy ước coi là mục tiêu tối ưu của bài toán min là $+\infty$ và của bài toán max là $-\infty$. Vậy ta có lỗ hổng đối ngẫu là rộng vô cùng.

(d) Lý thuyết đối ngẫu cho một phương tiện tuyệt vời để kiểm ra nghiệm tối ưu. Giả sử bạn phải giải một quy hoạch cỡ lớn rất khó để giúp bà giám đốc quyết định chiến lược cạnh tranh của công ti. Sau khi áp dụng phương pháp đơn hình và tìm được nghiệm tối ưu x^*, làm sao bạn thuyết phục được bà giám đốc tính đúng đắn của nó. Vì quyết định kinh tế dựa vào x^* là mang tính sống còn nên giám đốc cần sự đảm bảo. Kiểm tra lại cả quá trình tính toán chăng, đây là việc không tưởng! Bạn hãy áp dụng lý thuyết đối ngẫu. Theo chứng minh Định lí đối ngẫu mạnh 6.3.4, khi giải quy hoạch gốc bằng phương pháp đơn hình bạn cũng đồng thời có nghiệm tối ưu của bài toán đối ngẫu là y^*, với $y_i^* = -d_i^*$, $i = 1,..., m$. Bạn chỉ cần chỉ định lí đó cho bà giám đốc và kiểm tra cho bà ta thấy x^* là chấp nhận được (dễ), y^* là chấp nhận được cho bài toán đối ngẫu (cũng dễ như vậy) và $(y^*)^T b = c^T x^*$ (còn dễ hơn).

d_i = hệ số x biến bù

Các giáo sư của bạn cũng xem xét bài thi hoặc kiểm tra của bạn như vậy đó. Cho nên tốt nhất bạn hãy tự kiểm tra bài của mình theo cách đó trước.

(e) Lý thuyết đối ngẫu còn giúp xây dựng các thuật toán liên quan đến tính đối ngẫu mà trong nhiều trường hợp cho lời giải nhanh hơn phương pháp đơn hình thông thường. Thí dụ đơn giản như trường hợp chưa biết nghiệm cơ sở chấp nhận được của bài toán gốc nhưng lại biết nghiệm cơ sở chấp nhận được của bài toán đối ngẫu thì ta có thể thay cho việc làm cả hai pha đối với bài toán gốc bằng cách áp dụng chỉ pha II cho bài toán đối ngẫu.

6.4 ĐỘ LỆCH BÙ

Định lí đối ngẫu mạnh cho ta tiêu chuẩn để kiểm tra tính tối ưu như ở Nhận xét 6.3.1 (d). Nói cách khác nó cho ta điều kiện tối ưu: $c^Tx^* = (y^*)^Tb$. Điều kiện này biểu thị qua giá trị hàm mục tiêu. Từ định lí đối ngẫu cũng suy được ra điều kiện tối ưu ở dạng biểu thị qua các dữ kiện của bài toán. Điều kiện này thường gọi là *điều kiện độ lệch bù* (complementary slackness). Trong kinh tế định lí này thường có tên là *định lí cân bằng* (equilibrium theorem) vì trạng thái tối ưu có ý nghĩa là trạng thái cân bằng kinh tế.

ĐỊNH LÍ 6.4.1 (điều kiện độ lệch bù). Giả sử x và y là nghiệm chấp nhận được của bài toán gốc và bài toán đối ngẫu, tương ứng. Khi đó x và y là tối ưu khi và chỉ khi

$$y_i(a_i^Tx - b_i) = 0 \quad \forall i, \tag{10}$$

$$(c_j - y^TA_j)x_j = 0 \quad \forall j. \tag{11}$$

CHỨNG MINH. Ở chứng minh Định lí 6.3.1 ta đã đặt $u_i = y_i(a_i^Tx - b_i)$ và $v_j = (c_j - y^TA_j)x_j$ và thấy là theo định nghĩa bài toán đối ngẫu ta có $u_i \geq 0$, $v_j \geq 0$ $\forall i, \forall j$. Đồng thời ta cũng có

$$c^Tx - y^Tb = \sum_i u_i + \sum_j v_j .$$

Theo định lí đối ngẫu mạnh, nếu x và y là tối ưu thì $c^Tx = y^Tb$. Do đó $u_i = v_j = 0$ $\forall i, \forall j$. Ngược lại, nếu $u_i = v_j = 0$ $\forall i, \forall j$ thì $c^Tx = y^Tb$. Theo Hệ quả 6.3.3, x và y cũng là tối ưu. □

Các điều kiện nêu ở Định lí 6.4.1 gọi là điều kiện độ lệch bù vì từ đó ta có

$$a_i^T x < b_i \quad \Rightarrow \quad y_i = 0, \tag{12}$$

$$y^T A_j < c_j \quad \Rightarrow \quad x_j = 0. \tag{13}$$

(12) nghĩa là ràng buộc ở bài toán gốc "lệch" khỏi thỏa chặt thì biến đối ngẫu tương ứng "bù" lại phải không lệch khỏi thỏa chặt, tức là bằng 0 (và tương tự cho điều kiện (13)). Điều này rất tự nhiên nếu nói theo ngôn ngữ bài toán mở đầu ở Mục 6.1. Một ràng buộc không thỏa chặt tại điểm tối ưu thì có thể bỏ đi mà không ảnh hưởng đến giá trị mục tiêu tối ưu, và do đó không cần phải đặt giá y_i khác không cho sự vi phạm ràng buộc đó.

NHẬN XÉT 6.4.1

(a) Nếu bài toán gốc ở dạng chính tắc thì (10) luôn thỏa với mọi x chấp nhận được nên điều kiện độ lệch bù cho quy hoạch chính tắc chỉ còn là

$$(c_j - y^T A_j)x_j = 0 \quad \forall j.$$

(b) Điều kiện độ lệch bù có thể cho ta có ngay nghiệm tối ưu của quy hoạch tuyến tính khi biết nghiệm tối ưu của bài toán đối ngẫu. Hãy xét thí dụ sau

THÍ DỤ 6.4.1. Xét cặp bài toán đối ngẫu

$$\begin{array}{ll}
\min\ 13x_1 + 10x_2 + 6x_3, & \max\ 8y_1 + 3y_2, \\
\quad 5x_1 + x_2 + 3x_3 = 8, & \quad 5y_1 + 3y_2 \le 13, \\
\quad 3x_1 + x_2 \qquad\quad = 3, & \quad y_1 + y_2 \le 10, \\
\quad x_1, x_2, x_3 \ge 0. & \quad 3y_1 \qquad \le 6.
\end{array}$$

Ngay dưới đây ta sẽ kiểm tra là $x^* = (1, 0, 1)$ là nghiệm tối ưu của bài toán gốc. Trước mắt ta thấy nó chấp nhận được (và giả sử nó tối ưu) và sử dụng điều kiện độ lệch bù (11) (chỉ có (11) vì bài toán gốc có dạng chính tắc)

$$(c_j - y^T A_j)x_j^* = 0 \qquad j = 1, 2, 3.$$

Vì $x_2^* = 0$ nên điều kiện thỏa với j = 2. Còn $x_1^* > 0$ và $x_3^* > 0$ nên ta có, theo điều kiện độ lệch bù

$$
\begin{aligned}
5y_1 + 3y_2 &= 13,\\
3y_1 \qquad &= 6.
\end{aligned}
$$

Giải hệ này ta được $y_1^* = 2$ và $y_2^* = 1$. Tính giá trị mục tiêu đối ngẫu ta được $8y_1^* + 3y_2^* = 19$. Vì x^* là chấp nhận được, $y^* = (y_1^*, y_2^*)$ là tối ưu theo điều kiện độ lệch bù. Giá trị mục tiêu ứng với x^* cũng là 19 nên x^* cũng là tối ưu (của bài toán gốc).

Nhận xét rằng x^* là nghiệm tối ưu cơ sở không suy biến (vì có 2 = n - m thành phần dương). Trường hợp ta có một nghiệm tối ưu cơ sở suy biến thì điều kiện độ lệch bù giúp rất ít cho việc xác định nghiệm tối ưu cho bài toán đối ngẫu như người ta có thể chứng minh một cách tổng quát được.

6.5 THUẬT TOÁN ĐƠN HÌNH ĐỐI NGẪU

Ở Mục 6.3, khi chứng minh định lí đối ngẫu mạnh ta đã biết là giải một bài toán quy hoạch tuyến tính thì ta cũng được nghiệm của bài toán đối ngẫu. Ở nhận xét (e) cuối mục đó ta cũng đã chú ý là nếu bài toán gốc khó giải (chẳng hạn chưa có từ vựng chấp nhận được để xuất phát) ta có thể giải bằng cách áp dụng thuật toán đơn hình cho bài toán đối ngẫu. Tất nhiên do có đối ngẫu mạnh, cuối cùng ta vẫn được nghiệm tối ưu của bài toán gốc. Bây giờ ta xét cụ thể điều này qua một thí dụ

THÍ DỤ 6.5.1. Giả sử ta phải giải quy hoạch

$$
\begin{aligned}
&\min x_1 + x_2,\\
&-2x_1 - x_2 \le 4,\\
&-2x_1 + 4x_2 \le -8,
\end{aligned}
\qquad (14)
$$

$$-\ x_1 + 3x_2 \le -7,$$
$$x_1, x_2 \ge 0.$$

Vì trong các vế phải b_i có số âm nên khi viết thành từ vựng (sau khi đưa biến bù w_1, w_2, vào) ta được từ vựng cơ sở không chấp nhận được. Nếu muốn dùng đơn hình ta phải qua pha I. Để tránh việc đó ta hãy viết bài toán đối ngẫu của (14) và biến đổi nó về dạng chuẩn ở bên phải (y_i mới là $-y_i$ cũ)

$$\begin{array}{ll}
\max \zeta = 4y_1 - 8y_2 - 7y_3, & \min -\zeta = 4y_1 - 8y_2 - 7y_3,\\
y_1 \le 0, & 2y_1 + 2y_2 + y_3 \le 1,\\
y_2 \le 0, & y_1 - 4y_2 - 3y_3 \le 1,\\
y_3 \le 0, & y_1, y_2, y_3 \ge 0.\\
-2y_1 - 2y_2 - y_3 \le 1, & \\
-y_1 + 4y_2 + 3y_3 \le 1. &
\end{array} \qquad (15)$$

Dưới đây ta sẽ viết từ vựng và thực hiện thuật toán đơn hình. Ở cột bên trái luôn là quy hoạch gốc, cột bên phải là quy hoạch đối ngẫu. Trước hết ta viết cặp đối ngẫu (14) và (15) ở dạng từ vựng

$$\begin{array}{llll}
z = & & x_1 & + x_2\\
\hline
w_1 = 4 & & + 2x_1 & + x_2\\
w_2 = -8 & & + 2x_1 & - 4x_2\\
w_3 = -7 & & + x_1 & - 3x_2.
\end{array}
\qquad
\begin{array}{llll}
-\zeta = & 4y_1 & - 8y_2 & - 7y_3\\
\hline
z_1 = 1 & - 2y_1 & - 2y_2 & - y_3\\
z_2 = 1 & - y_1 & + 4y_2 & + 3y_3.
\end{array}$$

Nhận xét là từ vựng đối ngẫu nhận được bằng cách chuyển vị (tức là đổi hàng thành cột) và có thay đổi dấu từ vựng gốc. Việc thay dấu bao gồm giữ nguyên dấu ở hàng mục tiêu và các hằng số (tức là cột vế phải) và đổi dấu ở các hệ số khác (tức là các a_{ij}).

Vì từ vựng gốc không chấp nhận được nhưng từ vựng đối ngẫu chấp nhận được nên ta áp dụng thuật toán đơn hình (với quy tắc xoay Dantzig) lên từ vựng đối ngẫu. Tuy vậy ta vẫn đồng thời thực hiện phép toán tương ứng lên từ vựng gốc để theo dõi mối quan hệ. Dễ thấy biến vào là y_2 và biến ra là z_1. Vì w_2 (ứng với ràng buộc gốc thứ hai) là biến bù của y_2 và x_1 là biến bù của z_1, nên cặp w_2, x_1 sẽ là biến vào/ra tương ứng ở quy hoạch gốc. Nhưng x_1 đang ở ngoài cơ sở nên phải là biến vào (ngược với z_1) và w_2 đang ở trong cơ sở phải là biến ra (ngược với y_2).

Thực hiện phép xoay đã chọn cho cả hai từ vựng ta được

$$\begin{array}{l}
z = 4 + 0{,}5w_2 + 3x_2 \\
\hline
w_1 = 12 + w_2 + 5x_2 \\
x_1 = 4 + 0{,}5w_2 + 2x_2 \\
w_3 = -3 + 0{,}5w_2 - x_2.
\end{array}
\qquad
\begin{array}{l}
-\zeta = -4 + 12y_1 + 4z_1 - 3y_3 \\
\hline
y_2 = 0{,}5 - y_1 - 0{,}5z_1 - 0{,}5y_3 \\
z_2 = 3 - 5y_1 - 2z_1 + y_3
\end{array}$$

(Chú ý là hằng số hàng mục tiêu ở hai từ vựng là trái dấu.)

Thực hiện bước đơn hình tiếp theo trên từ vựng đối ngẫu ta thấy y_3 vào và y_2 ra. Tương ứng ở từ vựng gốc là w_2 vào và w_3 ra:

$$\begin{array}{l}
z = 7 + w_3 + 4x_2 \\
\hline
w_1 = 18 + 2w_3 + 7x_2 \\
x_1 = 7 + w_3 + 3x_2 \\
w_2 = 6 + 2w_3 + 2x_2.
\end{array}
\qquad
\begin{array}{l}
-\zeta = -7 + 18y_1 + 7z_1 + 6y_2 \\
\hline
y_3 = 1 - 2y_1 - z_1 - 2y_2 \\
z_2 = 4 - 7y_1 - 3z_1 - 2y_2.
\end{array}$$

Đến đây ta thấy cả hai từ vựng đều đã tối ưu. Chú ý rằng ở mỗi bước quy luật về chuyển vị và đổi dấu ở hai từ vựng đều như nhau.

Tất nhiên ta chỉ cần giải quy hoạch đối ngẫu rồi áp dụng định lí đối ngẫu mạnh cũng được ngay lời giải cho bài toán gốc: mục tiêu tối ưu là $z^* = \zeta^* = 7$, nghiệm tối ưu là $(x_1^*, x_2^*) = (d_1^*, d_2^*) = (7, 0)$. (Hãy so sánh với $y_i^* = -d_i^*$ ở chứng minh định lí đối ngẫu mạnh!) Nhưng bằng nhận xét mối quan hệ trên đây giữa từ vựng gốc và từ vựng đối ngẫu cũng như giữa hai phép xoay ta thấy có thể làm ngược lại là chỉ viết dãy từ vựng gốc nhưng lại thực hiện phép xoay tương ứng với phép xoay đơn hình ở từ vựng đối ngẫu (mà không cần viết dãy từ vựng đối ngẫu). Thuật toán này gọi là *thuật toán đơn hình đối ngẫu* (dual simplex algorithm).

THUẬT TOÁN ĐƠN HÌNH ĐỐI NGẪU

(thực hiện trên từ vựng gốc mỗi bước lặp như sau)

- Biến ra là biến có hằng số vế phải âm nhất (tức là âm và modul lớn nhất). Nếu không có biến như vậy (mà từ vựng đối ngẫu là chấp nhận được, tức là các hệ số hàng mục tiêu gốc đều không âm) thì từ vựng hiện hành đã là tối ưu.

- Để chọn biến vào ta xét tỉ số dương lớn nhất giữa các hệ số hàng ra đã chọn và hệ số tương ứng ở hàng mục tiêu. Biến ứng với tỉ số dương lớn nhất là biến vào.

- Phép xoay được thực hiện hoàn toàn như phép xoay đơn hình.

GIẢI THÍCH HÌNH HỌC THUẬT TOÁN ĐƠN HÌNH ĐỐI NGẪU

Ta đã định nghĩa thuật toán đơn hình đối ngẫu là thuật toán thực hiện trên quy hoạch gốc khi từ vựng xuất phát của nó không chấp nhận được, tương ứng với thuật toán đơn hình thực hiện trên quy hoạch đối ngẫu với từ vựng xuất phát chấp nhận được.

Ta cũng đã nhận xét là ở mỗi bước lặp, nếu ta viết cả hai từ vựng gốc và đối ngẫu ra thì thấy chúng có tính chuyển vị lẫn nhau với sự đổi dấu chỉ ở phần tử góc trên bên trái (giao của hàng mục tiêu và cột số hạng tự do) và ở các hệ số a_{ij}, còn ở hàng mục tiêu và cột số hạng tự do thì dấu không đổi (chỉ chuyển vị).

Về mặt hình học ta hãy xem thuật toán có ý nghĩa thế nào đối với bài toán gốc.

• Đỉnh xuất phát là đỉnh không chấp nhận được, tức là giao của các diện của tập lồi đa diện chấp nhận được, nhưng nằm ngoài tập này. Điều này rõ ràng vì từ vựng xuất phát không chấp nhận được, tức là nó xác định một nghiệm cơ sở không chấp nhận được.

• Qua mỗi bước lặp, đúng một biến vào cơ sở và một biến ra khỏi cơ sở. Việc này có nghĩa là đi từ một đỉnh sang đỉnh kề với nó như ở Mục 4.5 đã giải thích. Bước lặp tương ứng ở quy hoạch đối ngẫu làm mục tiêu ζ tăng ($-\zeta$ giảm) mà ζ luôn bằng giá trị mục tiêu z của bài toán gốc. Vậy qua mỗi bước lặp của thuật toán đơn hình đối ngẫu, mục tiêu gốc z tăng lên (chứ không cải thiện hơn). Tuy vậy, nói chung là đỉnh mới sẽ "gần" vào miền chấp nhận được hơn, vì các hệ số mục tiêu âm ở quy hoạch đối ngẫu "ít đi" (ta đã chọn hệ số âm nhất cho biến vào!) tức là các số âm trong cột số hạng tự do của từ vựng gốc "ít đi", đồng nghĩa với sự "giảm tính không chấp nhận được" của đỉnh tương ứng với từ vựng.

• Thuật toán ngừng khi từ vựng đối ngẫu tối ưu tức là hệ số hàng mục tiêu đều không âm (còn cột số hạng tự do luôn không âm ở mỗi bước lặp). Lúc đó từ vựng gốc có cột số hạng tự do lần đầu tiên trở thành không âm, tức là đỉnh chấp nhận được. Còn hệ số hàng mục tiêu của nó là cột số hạng tự do của từ vựng đối ngẫu nên đã luôn không âm qua mọi bước lặp. Vậy từ vựng gốc cũng là tối ưu.

Tóm lại, thuật toán đơn hình đối ngẫu xuất phát từ một đỉnh không chấp nhận được, nhưng tương ứng với một đỉnh chấp nhận được của quy hoạch đối ngẫu. Qua mỗi bước lặp, ta chuyển sang đỉnh kề "gần hơn" theo nghĩa tiến dần về miền chấp nhận được, với sự trả giá là mục tiêu kém đi (tức là tăng lên). Đỉnh đầu tiên vào được miền chấp nhận chính là đỉnh tối ưu. Ta thấy rõ sự đối lại với *thuật toán đơn hình gốc* (primal simplex algorithm), ở đây ta thêm chữ "gốc" để nhấn mạnh sự khác nhau. Ở thuật toán đơn hình gốc ta đi từ một đỉnh chấp nhận được, qua các đỉnh kề nhưng luôn luôn ở trong miền chấp nhận được, sao cho qua mỗi bước hàm mục tiêu giảm, cho đến đỉnh tối ưu thì dừng.

6.6 PHA MỘT DỰA TRÊN TÍNH ĐỐI NGẪU

Nhờ thuật toán đơn hình đối ngẫu, trường hợp từ vựng xuất phát chưa chấp nhận được nhưng từ vựng đối ngẫu tương ứng chấp nhận được, ta sẽ tránh được pha I bổ trợ mà tiến hành ngay thuật toán.

Nếu cả từ vựng đối ngẫu cũng không chấp nhận được thì sao? Ngay sau đây ta sẽ thấy rằng thuật toán đơn hình đối ngẫu cũng cho ta một pha I hay không kém gì pha I xét ở Mục 2.3. Ý tưởng đơn giản như sau. Ta cải biên bài toán gốc để đối ngẫu với bài toán cải biên có từ vựng xuất phát chấp nhận được. Muốn vậy, vì cột số hạng tự do (quyết định về tính chấp nhận được) của từ vựng đối ngẫu là các hệ số hàng mục tiêu gốc, ta chỉ việc cải biên hàm mục tiêu gốc để các hệ số đều không âm.

Đã có từ vựng đối ngẫu (cải biên) chấp nhận được rồi thì ta thực hiện được thuật toán đơn hình đối ngẫu trên bài toán cải biên và đi đến từ vựng tối ưu của bài toán cải biên. Vì bài toán cải biên có chung miền ràng buộc với bài toán ban đầu nên từ vựng cuối cùng này là chấp nhận được của bài toán ban đầu.

Lúc này ta có thể tiến hành pha II là thuật toán đơn hình gốc trên quy hoạch ban đầu. Thuật toán hai pha này được gọi là *thuật toán đối ngẫu - gốc* (dual - primal algorithm) hoặc thật đầy đủ là *thuật toán đơn hình hai pha đối ngẫu - gốc* (dual - primal two - phase simplex algorithm). Nhận xét thêm rằng pha I của thuật toán đối ngẫu - gốc cũng giúp ta phát hiện được bài toán gốc là không chấp nhận được. Thật vậy, nếu bài toán đối ngẫu của bài toán gốc cải biên là không giới nội thì bài toán gốc cải biên, và do đó bài toán gốc cũng vậy, là không chấp nhận được.

Ta cũng thấy rằng có thể thực hiện hai pha theo thứ tự ngược lại như sau để được thuật toán mới. Ta cải biên cột số hạng tự do của từ vựng gốc để từ vựng này trở thành chấp nhận được. Khi đó ta tiến hành được thuật toán đơn hình gốc để nhận được từ vựng tối ưu của bài toán cải biên. Từ vựng đối ngẫu tương ứng là từ vựng tối ưu của bài toán đối ngẫu của bài toán cải biên, nhưng nó vẫn bảo đảm là từ vựng chấp nhận được của bài toán đối ngẫu ban đầu (vì cải biên cột số hạng tự do ở từ vựng gốc xuất phát tương ứng là cải biên hàng mục tiêu của từ vựng đối ngẫu; riêng tính chấp nhận được không bị ảnh hưởng). Sau khi thực hiện pha I như vậy trên từ vựng gốc, pha II sẽ là thuật toán đơn hình trên từ vựng đối ngẫu chấp nhận được vừa có để đi đến từ vựng đối ngẫu tối ưu. Từ vựng tương ứng bên gốc sẽ là từ vựng tối ưu cần tìm. Thuật toán hai pha này được gọi là *thuật toán gốc - đối ngẫu* (primal - dual algorithm) hoặc có tên đầy đủ là *thuật toán đơn hình hai pha gốc - đối ngẫu* (primal - dual two - phase simplex algorithm).

BÀI TẬP CHƯƠNG 6

6.1. Tìm đối ngẫu của quy hoạch

$$\max c^T x,$$
$$a \le Ax \le b,$$
$$l \le x \le u.$$

6.2. Minh họa định lí đối ngẫu mạnh bằng kết quả của Bài tập 2.5 (sau khi đã giải bài tập này).

6.3. Xét quy hoạch tuyến tính chính tắc với A có các hàng độc lập tuyến tính. Chứng minh hoặc tìm phản thí dụ cho mỗi điều khẳng định sau đây.

(a) Giả sử x^* là nghiệm cơ sở chấp nhận được. Giả sử với mỗi cơ sở tương ứng x^*, nghiệm cơ sở tương ứng của bài toán đối ngẫu là không chấp nhận được. Khi đó giá trị mục tiêu tối ưu nhỏ hơn (hẳn) $c^T x^*$.

(b) Bài toán đối ngẫu của bài toán bổ trợ ở pha I của phương pháp đơn hình hai pha luôn chấp nhận được.

(c) Giả sử y_i là biến đối ngẫu tương ứng ràng buộc đẳng thức thứ i. Loại bỏ ràng buộc đẳng thức thứ i ở bài toán gốc là tương đương với đưa thêm vào ràng buộc $y_i = 0$ ở bài toán đối ngẫu.

6.4. Giải các quy hoạch ở Bài tập 2.6 bằng thuật toán đơn hình đối ngẫu.

6.5. Giải các quy hoạch ở Bài tập 2.6 bằng thuật toán đơn hình đối ngẫu-gốc.

6.6. Giải bằng thuật toán đơn hình đối ngẫu - gốc

(a)
$$\begin{aligned} &\max -x_1 - 3x_2 - x_3,\\ &2x_1 - 5x_2 + x_3 \le -5,\\ &2x_1 - x_2 + 2x_3 \le 4,\\ &x_1, x_2, x_3 \ge 0; \end{aligned}$$

(b)
$$\begin{aligned} &\min -7x_2 + 3x_3,\\ &-x_1 - 2x_2 - x_3 \le -3,\\ &-x_1 \le 1,\\ &-x_3 \le 0,\\ &x_1 + 3x_2 + 2x_3 = 27,\\ &2x_1 - x_2 + 3x_3 = 12; \end{aligned}$$

(c)
$$\begin{aligned} &\max 3x_1 + x_2,\\ &x_1 - x_2 \le -1,\\ &-x_1 - x_2 \le -3,\\ &2x_1 + x_2 \le 2,\\ &x_1, x_2 \ge 0. \end{aligned}$$

6.7. Giải Bài tập 6.6 bằng thuật toán đơn hình gốc – đối ngẫu.

6.8. Giải

(a)
$$\begin{aligned} &\min 23x_1 - 7x_2,\\ &-4x_1 + x_2 \le -2,\\ &x_1 + x_2 \le 5,\\ &-x_1 - x_2 \le -1,\\ &-3x_1 + 2x_2 \le 1,\\ &x_1, x_2 \ge 0; \end{aligned}$$

(b)
$$\begin{aligned} \min\ & -9x_1 - x_2, \\ & 6x_1 + 5x_2 \le 10, \\ & 3x_1 - 2x_2 \le 8, \\ & 2x_1 + x_2 \le 3, \\ & -x_1 + 4x_2 \le -4; \end{aligned}$$

(c)
$$\begin{aligned} \min\ & -6x_1 - 4x_2 - 8x_3, \\ & x_1 + x_2 + x_3 \le 4, \\ & x_1 + 2x_2 - x_3 \le -1, \\ & 2x_1 + 3x_3 \le 8, \\ & 2x_2 + 4x_3 \le 10, \\ & x_1 + 2x_2 + 3x_3 \le 7; \end{aligned}$$

(d)
$$\begin{aligned} \max\ & x_1 + 3x_2 - x_3, \\ & 2x_1 + 2x_2 - x_3 \le 10, \\ & 3x_1 - 2x_2 + x_3 \le 10, \\ & x_1 - 3x_2 + x_3 \le 10, \\ & x_1, x_2, x_3 \ge 0. \end{aligned}$$

(e)
$$\begin{aligned} \min\ & 3x_1 - 2x_3 + x_4, \\ & x_1 + 2x_2 - 2x_3 - x_4 + x_5 \ge 5, \\ & 5x_1 + 6x_2 - 3x_3 - x_4 - x_5 \le 10, \\ & 3x_1 + 2x_2 + x_3 + x_4 - 3x_5 \ge 3, \\ & x_1, \ldots, x_5 \ge 0. \end{aligned}$$

6.9. Dùng phương pháp đơn hình mô tả *tất cả* các nghiệm tối ưu của quy hoạch tuyến tính

$$\begin{aligned} \max\ & 2x_1 + 3x_2 + 5x_3 + 4x_4, \\ & x_1 + 2x_2 + 3x_3 + x_4 \le 5, \\ & x_1 + x_2 + 2x_3 + 3x_4 \le 3, \\ & x_1, x_2, x_3, x_4 \ge 0. \end{aligned}$$

6.10. Xét quy hoạch

$$\max \sum_j p_j x_j,$$

$$\sum_j q_j x_j \le \beta,$$

$$x_j \le 1,\ j = 1,\dots, n,$$

$$x_j \ge 0,\ j = 1,\dots, n.$$

Giả sử các p_j, q_j dương và $\sum_j p_i = \sum_j q_i = 1$. Giả sử

$$\frac{p_1}{q_1} < \frac{p_2}{q_2} < \dots < \frac{p_n}{q_n}$$

và $\beta > 0$ đủ nhỏ. Kí hiệu

$$k = \min\{j: q_{j+1} + \dots + q_n \le \beta\},$$

y_0 là biến đối ngẫu tương ứng ràng buộc chứa β và y_j là biến đối ngẫu tương ứng ràng buộc $x_j \le 1$.

Dùng lí thuyết đối ngẫu chứng minh rằng nghiệm tối ưu của bài toán gốc và đối ngẫu tương ứng là

$$x_j = \begin{cases} 0 & j < k, \\ \dfrac{\beta - q_{k+1} - \dots - q_n}{q_k} & j = k, \\ 1 & j > k; \end{cases}$$

$$y_j = \begin{cases} \dfrac{p_k}{q_k} & j = 0, \\ 0 & 0 < j \le k, \\ q_j\left(\dfrac{p_j}{q_j} - \dfrac{p_k}{q_k}\right) & j > k. \end{cases}$$

(Bài này có nguồn gốc từ Bài tập 1.3.)

6.11. Giả sử A là ma trận vuông đối xứng. Xét quy hoạch

$$\min c^T x,\ Ax \geq c,\ x \geq 0.$$

Chứng minh rằng nếu x^* thỏa $Ax^* = c$, $x^* \geq 0$ thì x^* là nghiệm tối ưu.

6.12. Tìm bài toán đối ngẫu của quy hoạch

$$\begin{aligned}
\min\ & -6x_1 - 4x_2 - 8x_3,\\
& x_1 + x_2 + x_3 \leq 4,\\
& x_1 + 2x_2 - x_3 \leq -1,\\
& 2x_1 + 3x_3 \leq 8,\\
& 2x_2 + 4x_3 \leq 10,\\
& x_1 + 2x_2 + 3x_3 \leq 7.
\end{aligned}$$

Giả sử $x^0 = (1, 0, 2)^T$.

(a) Dùng tính chất độ lệch bù tìm nghiệm đối ngẫu tương ứng u^0.

(b) Dùng Định lí đối ngẫu yếu chứng tỏ rằng x^0 và u^0 là nghiệm tối ưu của hai bài toán gốc và đối ngẫu tương ứng.

(c) x^0 có phải là nghiệm tối ưu duy nhất không? Tại sao? Cũng câu hỏi như vậy cho u^0.

6.13. Tìm quy hoạch đối ngẫu của

$$\begin{aligned}
\min\ & -2u_1 - 14u_2 - 36u_3,\\
& -2u_1 + u_2 + 4u_3 - u_4 = 5,\\
& u_1 + 2u_2 + 3u_3 - u_5 = -2,\\
& -1 \leq u_i \leq 3,\ i = 1, \dots, 5.
\end{aligned}$$

6.14. Xét quy hoạch

$$\begin{aligned}
\min\ & -4x_1 - 5x_2 - 4x_3,\\
& x_1 + x_3 \leq 4,\\
& 2x_1 + x_2 - x_3 \leq 3,\\
& x_2 + x_3 \leq 5,\\
& -x_2 + x_3 \leq 2,
\end{aligned}$$

$$-x_1 + x_2 - x_3 \le -2,$$
$$x_1 + x_2 + x_3 \le 7.$$

Giả sử $x^0 = (1, 2, 3)^T$.

(a) x^0 có phải là điểm cực biên của tập ràng buộc không?

(b) Chứng minh rằng x^0 là nghiệm tối ưu.

(c) Viết bài toán đối ngẫu và chỉ ra một nghiệm tối ưu.

(d) Xác định tất cả các nghiệm tối ưu của bài toán gốc.

6.15. Xét quy hoạch min c^Tx, $Ax \le b$. Ta đã biết điều kiện độ lệch bù là $y_i(a_i^Tx - b_i) = 0$, $i = 1,\dots, m$, ở đây y_i là biến đối ngẫu. Điểm x^0 được gọi là thỏa *điều kiện độ lệch bù chặt* (strict complementary slackness condition) nếu x^0 thỏa điều kiện độ lệch bù và

$$\text{nếu } a_i^Tx^0 = b_i, \text{ thì } y_i > 0.$$

Giả sử x^0 là nghiệm cơ sở tối ưu không suy biến và thỏa điều kiện độ lệch bù chặt. Chứng minh rằng x^0 là nghiệm tối ưu duy nhất.

6.16. Xét quy hoạch min c^Tx, $A_1x \le b_1$, $A_2x = b_2$. Ta nói điểm x^0 hội thỏa *điều kiện độ lệch bù chặt* cho quy hoạch này nếu nó thỏa điều kiện độ lệch bù, và biến đối ngẫu tương ứng ràng buộc bất đẳng thức thỏa chặt là dương (hẳn). Chứng minh khẳng định về tính chất duy nhất nghiệm tối ưu như Bài tập 6.15.

6.17. Xét quy hoạch gốc min c^Tx, $Ax \ge b$, $x \ge 0$. Hãy viết quy hoạch đối ngẫu ở dạng bài toán tìm cực tiểu. Tìm điều kiện về A, b, c để bài toán đối ngẫu trùng với bài toán gốc. Cho thí dụ cụ thể thỏa điều kiện này.

6.18. Cho thí dụ một cặp quy hoạch tuyến tính đối ngẫu nhau mà cả hai quy hoạch đều có nhiều nghiệm tối ưu.

6.19. Giả sử một quy hoạch tuyến tính chính tắc không chấp nhận được nhưng khi bỏ đi ràng buộc đẳng thức cuối cùng thì trở thành chấp nhận được và có giá trị mục tiêu tối ưu hữu hạn. Chứng minh rằng quy hoạch đối ngẫu của quy hoạch gốc ban đầu là chấp nhận được và có giá trị mục tiêu tối ưu vô hạn.

6.20. Xét một quy hoạch tuyến tính tổng quát. Giả sử nó có nghiệm cơ sở chấp nhận được không suy biến. Chứng minh rằng điều kiện độ lệch bù sẽ là một hệ phương trình, của biến đối ngẫu, có nghiệm duy nhất.

6.21. Xét cặp quy hoạch tuyến tính chính tắc và đối ngẫu của nó.

(a) Chứng minh rằng nếu một trong hai quy hoạch có nghiệm tối ưu duy nhất và không suy biến thì quy hoạch kia cũng vậy.

(b) Giả sử ta có một cơ sở tối ưu không suy biến của quy hoạch gốc và biết rằng một hệ số mục tiêu của biến cơ sở là 0. Khi đó khẳng định rằng (a) suy ra điều gì? Có cơ sở tối ưu khác hay không?

6.22. Cho thí dụ trường hợp quy hoạch gốc có nghiệm cơ sở tối ưu suy biến nhưng quy hoạch đối ngẫu có nghiệm tối ưu duy nhất. (Không cần thí dụ ở dạng chính tắc.)

6.23. Xét quy hoạch min x_2, $x_2 = 1$, $x_1 \geq 0$, $x_2 \geq 0$. Hãy lập quy hoạch đối ngẫu. Xét xem hai quy hoạch này có nghiệm tối ưu duy nhất không? có nghiệm tối ưu không suy biến không? Thí dụ này có phù hợp với khẳng định rằng sự không suy biến của một nghiệm tối ưu cơ sở ở một bài toán suy ra sự duy nhất nghiệm tối ưu ở bài toán kia hay không?