Corso:
Ricerca Operativa I (co-docenza)
LS Ingegneria dei Trasporti e della Logistica, Università di Genova, 2006/07

Fioravante Patrone
Sezione Metodi e Modelli Matematici, DIPTEM
Facoltà di Ingegneria
Università di Genova

Solo i colti amano imparare; gli ignoranti preferiscono insegnare - Edouard Le Berquier

Lu 18 settembre 2006, 1h-1h:
Introduzione al corso.
DOCUMENTI:
.
ESERCIZI:
.

Gi 21 settembre 2006, 2h-3h:
Massimi e minimi, definizioni.
Parliamo di min (per max basta cambiare segno)
min per $f : A \to B$ con A e B insiemi qualsiasi non ha senso
$B = \mathbb{R}$. $f : A \to \mathbb{R}$.
Def di punto di min globale e di (valore) minimo globale (dedotta da quella di punto di min).
Si può anche seguire strada opposta, partendo da min di $f(A)$.
Punto di min stretto.
Non ha senso parlare di min locale. Dobbiamo poter parlare di intorni. Prendiamo $A \subseteq \mathbb{R}^n$
$f : A \subseteq \mathbb{R}^n \to \mathbb{R}$.
Def di min locale: c'è intorno V (sferico, se vogliamo) di $\bar{x}$ t.c. $\bar{x}$ sia punto di min (globale) per $f|V$

Problemi di min vincolato. Non è netta la distinzione fra pb libero e vincolato. Un pb libero lo possiamo vedere come vincolato e viceversa. Dipende da come ci sono forniti i dati, ed anche da quali siano le tecniche più opportune da applicare, nello specifico contesto in cui ci si trova.

Vincoli di uguaglianza: $h_i : A \to \mathbb{R}$, $i = 0, 1, \ldots, m$. Se $m = 0$ non abbiamo vincoli di uguaglianza...
Vincoli di disuguaglianza: $g_j : A \to \mathbb{R}$, $j = 0, 1, \ldots, r$. Se $r = 0$ non abbiamo vincoli di disuguaglianza...
Un po' strano che tutte le funzioni siano definite sullo stesso A, ma non è condizione restrittiva.

CN di min libero per funzioni di una variabile.
Proviamo a ricostruire il teorema.
Che f sia derivabile nel punto di min loc non lo possiamo ottenere come tesi, vedi esempio di $|x|$.
Proviamo a dimostrare il teorema. E' data $f : A \to \mathbb{R}$, con $A \subseteq \mathbb{R}^n$. E $\bar{x}$ è p.to di min loc.
Per la dim, tre idee:
1. considerare separatamente $\bar{x} < x$ e $\bar{x} > x$.
2. scrivere il rapporto incrementale e notare che ha segno definito (è $ge0$ se $\bar{x} < x$) .
3. usare teorema di permanenza del segno
Si ha allora $f'(\bar{x}) \geq 0$, arrivando da destra. Se si arriva da sinistra, $f'(\bar{x}) \leq 0$.
Allora $f'(\bar{x}) = 0$. Ma attenzione, stiamo dimenticando cose importantissime!!!
DOCUMENTI:
.
ESERCIZI:
.

Ve 22 settembre 2006, 2h-5h:
Finita la revisione della condizione $f'(x) = 0$: limiti da destra e da sinistra
Per arrivare a $f'(\bar{x}) = 0$, aggiungiamo la condizione che $\bar{x}$ sia interno ad A (la condizione davvero necessaria è che $\bar{x}$ sia punto di accumulazione sia da destra che da sinistra).
Sono comunque interessanti anche le condizioni unilaterali che, nel caso di un intervallo, si hanno negli estremi (se $A = [a, b]$, ed a è minimo, $f'(a) \geq 0$; se b è minimo, $f'(b) \leq 0$.
Generalizzazione ad $\mathbb{R}^n$: sia $\bar{x}$ un p.to di min locale. Allora, dato un vettore $v \in \mathbb{R}^n$, con $v \neq 0$, $\frac{\partial f}{\partial v}(\bar{x}) \geq 0$ se ha senso. [Con $\frac{\partial f}{\partial v}(x)$ indichiamo $\lim_{t \to 0} \frac{f(x+tv)-f(x)}{t}$] .
In particolare, se f è parzialmente derivabile ed $\bar{x}$ è interno ad A, per ogni $i \in \{1, \ldots, n\}$ si ha $\frac{\partial f}{\partial x_i}(\bar{x}) = 0$.

Per lavorare tranquillamente useremo delle ipotesi più forti su f, tipo la differenziabilità.
Ma la verifica diretta della differenziabilità può essere ostica. Se applicabile, possiamo sfruttare il fatto che, se $f \in \mathcal{C}^1(A)$, allora è differenziabile in ogni punto di A
Attenzione, però, che questo risultato richiede che l'insieme A sia aperto (la definizione di $\mathcal{C}^1(A)$, se A non è aperto, è molto laboriosa).

Un commento: quando le condizioni del primo ordine sono sufficienti? Per le funzioni convesse!
Data $f \in \mathcal{C}^1(A)$, con A aperto e convesso di $\mathbb{R}^n$, se f è convessa allora:
$\nabla f(\bar{x}) = 0$ se e solo se $\bar{x}$ è punto di minimo assoluto per f.
Lo si dimostra usando la caratterizzazione differenziale della convessità: $f \in \mathcal{C}^1(A)$, con A aperto, è convessa se e solo se per ogni $\bar{x}, x \in A$ si ha $f(x) \geq f(\bar{x}) + \nabla f(\bar{x}) \cdot (x - \bar{x})$.

Ci sono poi le condizioni del secondo ordine.
L'idea generale è questa. Ho f. La approssimo con una funzione semplice (ad esempio un polinomio). Guardo sotto quali condizioni questa funzione approssimante ha minimo. Cerco di provare se queste condizioni possano essere necessarie perché la f abbia minimo (o sufficienti, magari rafforzandole un poco: ad esempio, chiedendo che la funzione approssimante abbia un punto di minimo *stretto*
Supponiamo che sia $f \in \mathcal{C}^2(A)$
Noi approssimeremo f con il polinomio di Taylor (del secondo ordine): $f(x) + \sum_{i=1,\dots,n} \frac{\partial f}{\partial x_i}(\bar{x})(x_i - \bar{x}_i) + \sum_{i,j=1,\dots,n} \frac{\partial^2 f}{\partial x_i x_j}(\bar{x})(x_i - \bar{x}_i)(x_j - \bar{x}_j)$.
La parte su cui concentrare l'attenzione è la forma quadratica: $\sum_{i,j=1,\dots,n} \frac{\partial^2 f}{\partial x_i x_j}(\bar{x})(x_i - \bar{x}_i)(x_j - \bar{x}_j)$. Essa ha minimo in $x = \bar{x}$
Per comodità chiamiamo $\Phi(h) = \sum_{i,j=1,\dots,n} \frac{\partial^2 f}{\partial x_i x_j}(\bar{x}) h_i h_j$ e vediamo se Φ ha minimo in 0 .
Chiaramente Φ ha minimo in 0 se e solo se $\Phi(h) \geq 0$ per ogni $h \in \mathbb{R}^n$.
se $\Phi(h) \geq 0$ per ogni $h \in \mathbb{R}^n$, si dice che Φ è semidefinita positiva..
Come verificare se Φ è semidefinita positiva? Con la matrice hessiana: $H_{ij} = \frac{\partial^2 f}{\partial x_i x_j}(\bar{x})$. E' una matrice simmetrica ($f \in \mathcal{C}^2(A)$ e quindi vale il teorema di Schwartz). Quindi diagonalizzabile. Quindi Φ sarà semidefinita positiva se e solo se tutti i suoi autovalori saranno maggiori o uguali a zero.
Invece, la condizione $\Phi(h) > 0$ per ogni $h \neq 0$, cioè se Φ abbia in 0 un punto di minimo stretto, è una condizione sufficiente perché $\bar{x}$ sia un punto di minimo locale (ammesso che sia anche $\nabla f(\bar{x}) = 0$).
Se $\Phi(h) > 0$ per ogni $h \neq 0$, diciamo che Φ è definita positiva.

E Φ è definita positiva se e solo se ogni suo autovalore è strettamente positivo
Notare che se Φ ha un autovalore strettamente positivo ed uno strettamente negativo, il punto $\bar{x}$ non potrà essere né punto di minimo né punto di massimo
DOCUMENTI:
.
ESERCIZI:
.

Gi 28 settembre 2006, 3h-8h:
Lezione di Bagnerini.
DOCUMENTI:
Vedi i file: introduzione e algoritmi dal sito del corso di laurea in Ing Gest, Fac. di Ing., Univ. di Roma La Sapienza.
Vedi anche il libro di Maffioli (per dettagli, vedi bibliografia).
ESERCIZI:
.

Ve 29 settembre 2006, 2h-10h:
Il teorema di Weierstrass garantisce esistenza di p.to di minimo (globale).
Ipotesi: $f : A \to \mathbb{R}$, $A \subseteq \mathbb{R}^n$ chiuso e limitato, f continua.
C'è anche teorema di Weierstrass generalizzato: se A non è limitato, basta che $\lim_{||x|| \to \infty} f(x) = +\infty$.

Massimi e minimi vincolati (al solito, parleremo solo di min).
Abbiamo già visto il setting. Abbiamo $A \subseteq \mathbb{R}^n$ e sono date f,h_i, con $i = 0, \dots, m$ e g_j, $j = 0, \dots, r$. Tutte definite su A e a valori in $\mathbb{R}$
Cerchiamo p.ti di min (locale, globale) per $f|E$ dove E è l'insieme dei punti che soddisfano il vincolo (o, brevemente, il vincolo). Cioè è l'insieme dei punti $x \in \mathbb{R}^n$ t.c. $h_i(x) = 0$ e $g_j(x) \leq 0$ per ogni i e j.
Supponiamo $i \geq 1$ ed $r = 0$. Cioè, abbiamo vincoli di uguaglianza e solo quelli.
Supponiamo f e le h_i di classe $\mathcal{C}^1(A)$. Ricordare che serve A aperto, per parlare di $\mathcal{C}^1(A)$
Se $x^\star$ è punto di minimo vincolato, abbiamo la condizione detta dei moltiplicatori di Lagrange.
Occorre assumere anche che i vettori $\nabla h_i(\bar{x})$ siano *linearmente indipendenti.*
NOTA FUORI TESTO. Notare che questo implica $m \leq n$. A dire il vero, i caso $m = n$ è normalmente poco interessante: ci aspettiamo che il sistema

dei vincoli abbia un numero finito di soluzioni, ovvero di punti in $\mathbb{R}^n$ e quindi il problema di massimo diventa relativamente banale: A dire il vero, questi discorsi andrebbero molto precisati. Comunque, di solito si assume che il numero dei vincoli sia strettamente minore del numero delle incognite. FINE NOTA FUORI TESTO.
Allora esiste (ed è unico) un vettore $\lambda^\star = (\lambda_1^\star, \ldots, \lambda_m^\star) \in \mathbb{R}^m$ tale che:
$\nabla f(x^\star) + \sum_{i=1}^m \lambda_i^\star \nabla h_i(x^\star) = 0$.
Naturalmente $x^s tar$ soddisfa anche la condizione $h(x^s tar) = 0$ (cioè: $h_i(x^s tar) = 0$ per ogni i) .

Possiamo introdurre la lagrangiana: $L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i h_i(x)$.
Le condizioni le possiamo riscrivere allora così:.
- $\nabla_x L(x^\star, \lambda^\star) = 0$
- $\nabla_\lambda L(x^\star, \lambda^\star) = 0$

Fatto esempio. min $x + 2y$ con vincolo $x^2 + y^2 = 1$.
Osservato che la condizione sulla lineare indipendenza dei vettori $\nabla h_i(\bar{x})$ (qui ne abbiamo solo uno, quindi è come dire che questo vettore sia non nullo), va verificata per i punti che *stanno sul vincolo.*

Interpretazione geometrica (con due vincoli, in $\mathbb{R}^3$:
- il ∇f sta nello spazio vettoriale generato da ∇h_1 e ∇h_2
- il ∇f è ortogonale allo spazio vettoriale approssimazione lineare dei vincoli (Bertsekas lo chiama sottospazio delle variazioni ammissibili del primo ordine)

La dimostrazione (meglio, una delle possibili dimostrazioni) utilizza il teorema delle funzioni implicite (teorema di Dini). Lo si può dimostrare anche usando una idea, che è usata per la soluzione numerica di problemi di minimo vincolato: il metodo di penalizzazione.

Senza la condizione di lineare indipendenza, il teorema è falso. Prendere: $f(x_1, x_2) = x_1 + x_2$ (ad esempio, ma c'è ampia scelta...) e i vincoli $h_1(x_1, x_2) = (x_1 - 1)^2 + x_2^2 - 1$ e $h_2(x_1, x_2) = (x_1 - 2)^2 + x_2^2 - 4$. Il vincolo descrive un singleton. Ma la condizione di Lagrange:
$(1, 1) + \lambda_1(-2, 0) + \lambda_2(-4, 0) = (0, 0)$ non ha soluzioni

Se abbiamo anche vincoli di disuguaglianza, ci sono le condizioni di Kuhn-Tucker.
Stavolta il vincolo è dato da $h_i(x) = 0$ e $g_i(x) \leq 0$
Serve l'idea di vincolo *attivo*. Un vincolo di disuguaglianza (il vincolo j) è attivo in un punto $\bar{x}$ se $g_j(\bar{x}) = 0$. Indichiamo con $A(\bar{x})$ l'insieme degli *indici*

che identificano i vincoli attivi nel punto $\bar{x}$
Allora, se (oltre alle altre ovvie condizioni) la famiglia dei vettori:
- $\nabla h_1(x^\star), \ldots, \nabla h_m(x^\star)$
- $\nabla g_j(x^\star)$ con $j \in A(x^\star)$

è una famiglia di vettori linearmente indipendenti,
esiste ed è unico $(\lambda^\star, \mu^\star) = (\lambda_1^\star, \ldots, \lambda_m^\star, \mu_1^\star, \ldots, \mu_r^\star)$ t.c.:
- $\nabla_x L(x^\star, \lambda^\star, \mu^\star) = 0$
- $\mu_j^\star \geq 0$ per ogni j
- $\mu_j^\star = 0$ se $j \in A(x^\star)$

Esempio. svolto: min $x + 2y + 3z$ con vincoli: $x^2 + y^2 + z^2 \leq 4$ e $z \geq 1$.
Analizzati i 4 diversi casi che si possono presentare, a seconda di chi siano i vincoli attivi.
DOCUMENTI:
.
ESERCIZI:
.

Gi 5 ottobre 2006, 3h-13h:
Programmazione lineare. In programmazione matematica avevamo $A \subseteq \mathbb{R}^n$ ed f,h_i, con $i = 0, \ldots, m$ e g_j, $j = 0, \ldots, r$. Tutte definite su A e a valori in $\mathbb{R}$
Cercavamo p.ti di min (locale, globale) per $f|E$ dove E è l'insieme dei punti che soddisfano il vincolo (o, brevemente, il vincolo). Cioè è l'insieme dei punti $x \in \mathbb{R}^n$ t.c. $h_i(x) = 0$ e $g_j(x) \leq 0$ per ogni i e j.
Cosa cambia? Intanto, $A = \mathbb{R}^n$. E $f(x)$ è una funzione affine (ma possiamo assumerla lineare). Le funzioni h_i e g_j sono anch'esse affini.
Ci occuperemo di problemi di max anziché di minimo. Ma questo è solo una questione di abitudine nei due diversi contesti. Passare da un pb. di min a uno di max e viceversa comporta solo un cambiamento di segno.

Un esempio che sfrutteremo fino all'osso.
. $\max x_1 + x_2$
$6x_1 + 4x_2 \leq 24$
$3x_1 - 2x_2 \leq 6$
$x_1, x_2 \geq 0$

Notiamo che l'insieme dei vincoli risulta essere chiuso e convesso. Anzi, è un *poliedro* convesso (come vedremo).
Ma abbiamo da massimizzare una funzione concava (il che è come minimiz-

zare una funzione convessa, basta cambiare segno) su un insieme convesso. Allora i p.ti di max locale sono anche di max globale.
Ecco perché ci interessiamo dei max globali e solo di quelli: occuparsi di max locali sarebbe lo stesso...
Anzi, abbiamo un poliedro convesso (le definizioni precise la prossima volta). L'intuizione ci porta a ritenere che, se c'è massimo, questo sarà assunto in almeno un vertice. Occorre stare attenti: un poliedro potrebbe non avere vertici: striscia di piano fre due rette parallele. Ma la condizione che imporremo (tutte le variabili non-negative) ci consentirà di poterlo affermare.

Allora si potrebbe pensare di trovare tutti i vertici e tra questi selezionare quello (o quelli) dove la funzione obiettivo assume il valore massimo. Ma ci sono metodi migliori (da molti punti di vista).
Tra questi, il metodo del simplesso. Si parte da un vertice e ci si sposta da quello verso un vertice adiacente in modo da aumentare il valore della funzione obiettivo (vi saranno dei criteri che ci dicono se si può, e se la funzione può essere aumentata a piacere, essendo superiormente illimitata).

Una caratteristica specifica dei problemi di programmazione lineare è che, se la funzione è superiormente limitata sul vincolo, allora ha massimo. Questo non succede necessariamente per problemi di massimo non lineari: $-\exp(x)$, oppure $\mathrm{arctg}(x)$ sono esempi in tal senso.

Visti 3 esempi.
- problema della dieta
- modello di produzione lineare
- ricerca del flusso massimo in una rete

Ogni problema di PL può essere messo in forma canonica. (Visto come trattare le variabili senza vincolo di segno)
$\max cx$
$Ax \leq b$
$x \geq 0$
Ogni problema di PL può essere messo in forma standard.
$\max cx$
$Ax = b$
$x \geq 0$
Significato di equivalenza dei problemi.

Noi lavoreremo con pb. in forma standard. Anzi, lavoreremo nelle ipotesi:

- $b_j \geq 0$ per ogni $j \in \{1, \ldots, m\}$.
- $m < n$
- $r(A) = m$

La prima condizione la possiamo soddisfare facilmente (con un cambiamento di segno)
Per le altre due condizioni, se la seconda non è soddisfatta lo vediamo facilmente.
La terza condizione non è di agevole verifica. Tuttavia, il metodo del simplesso ci permetterà, tra le altre cose, di verificare se è soddisfatta oppure no.
DOCUMENTI:
.
ESERCIZI:
.

Ve 6 ottobre 2006, 2h-15h:
Variabili di slack e di surplus. E ricordato come si trattano le variabili libere (cioè senza vincolo di segno).

Interpretazione geometrica.
Supponiamo, per comodità, di avere un problema di PL in forma canonica.
Chiamiamo P l'insieme dei punti di $\mathbb{R}^n$ che soddisfano il vincolo.
Lo possiamo vedere come intersezione di un numero finito di semispazi.
Quindi è convesso (i semispazi sono convessi e la intersezione di convessi è convessa).
E' un convesso speciale. Non tutti i convessi sono intersezione di un numero finito di semispazi.
Visto che i convessi che sono intersezione di un numero finito di semispazi sono detti poliedri, P è un poliedro.
Un poliedro può essere limitato (in tal caso lo chiamiamo politopo) o illimitato. Se è limitato, può essere vuoto oppure no,
Ci interessano i punti estremi di P. Li chiameremo vertici. Di fatto, corrispondono all'idea elementare di vertice.
Teorema di Motzkin Ogni poliedro P può essere decomposto così: $P = \text{CONV}(X) + \text{CONE}(Y)$, con X ed Y finiti (Y può essere vuoto; lo è se e solo se P è limitato).
Se P è non vuoto ed inoltre $P \subseteq \mathbb{R}^n_{\geq}$ e X è minimale, allora P ha vertici e i suoi vertici sono gli elementi di X.
Se P è non limitato, come detto Y è non vuoto. Allora, cx ha massimo su P se e solo se $cy \leq 0$ per ogni $y \in Y$.

Interpretazione algebrica.
Supponiamo, per comodità, di avere un problema di PL in forma standard.
E anche che siano soddisfatte le condizioni: $b \geq 0$, $m < n$ e $r(A) = m$.
Prendiamo una sottomatrice B $m * m$ di A, che sia non-singolare.
La chiamiamo matrice di base (le sue colonne, e le variabili corrispondenti alle sue colonne le chiamiamo rispettivamente colonne e variabili di base).
Il sistema $Bx_B = b - F\bar{x}_F$ ha soluzione unica comunque fissiamo $\bar{x}_F$.
Prima osservazione.
Se fissiamo $\bar{x}_F = 0$, la soluzione che otteniamo è detta soluzione di base.
Non è detto che la soluzione di base $x_B = B^{-1}b$ soddisfi i vincoli di non negatività. Cioè non è detto che sia $B^{-1}b \geq 0$. Se questo è verificato, diremo che la soluzione è una soluzione ammissibile di base.
Si può dimostrare che, per un problema in forma standard, x è un vertice di P se e solo se è una soluzione ammissibile di base.
Nota: possono esserci più basi che danno luogo alla stessa soluzione ammissibile di base (e quindi allo stesso vertice).
Seconda osservazione.
Data una base B, possiamo esprimere le variabili di base in funzione delle variabili fuori base. Andando a sostituire nella espressione della funzione obiettivo, troviamo una funzione lineare nelle sole variabili fuori base. (O, se si vuole, anche nelle variabili di base, ma con coefficienti nulli rispetto a queste). I coefficienti di questa nuova funzione lineare trovata li chiamiamo costi ridotti (la terminologia sarebbe appropriata ad un problema di minimo, a dire il vero...).
Si prova facilmente che una soluzione di base, che sia ammissibile, è soluzione ottima del problema se e solo se i coefficienti di costo ridotti sono tutti minori o uguali di zero.
DOCUMENTI:

.
ESERCIZI:

.

Gi 12 ottobre 2006, 3h-18h:
Perché è naturale vedere c come vettore riga, ovvero come matrice $1 * n$ (è la matrice di rappresentazione di una applicazione lineare da $\mathbb{R}^n$ in R).

Teorema fondamentale della PL Sia dato un problema di PL in forma standard. Indico con P il poliedro dei vincoli. Allora:
- se $P! = \emptyset$, allora esiste una soluzione ammissibile di base (che, ricordo,

corrisponde a un vertice di P)
- se esiste una soluzione ottima, allora esiste una soluzione ammissibile di base ottima.
Siamo allora sicuri che, se il problema di PL ha soluzione, tra le soluzioni ci sarà almeno una soluzione ottima di base, ovvero un vertice di P.
Nota: tutto questo vale se abbiamo un problema in forma standard (andrebbe bene anche se fosse in forma canonica). Se non è in forma standard, allora non è detto che P abbia vertici. Basta pensare a: $\max x_1$ con $2 \leq x_1 \leq 5$ visto come problema in $\mathbb{R}^2$. (Se uno vuole un problema seriamente in $\mathbb{R}^2$, basta prendere $\max x_1 + x_2$ con vincoli $2 \leq x_1 + x_2 \leq 5$. Così è contento...).
Esercizio. Trasformare il problema dato in uno in forma standard ad esso equivalente e verificare che quest'ultimo ha una soluzione in un vertice. Trovare la soluzione corrispondente a questo vertice nel problema originario.

Occupiamoci ora di un problema di PL in forma standard, e supponiamo che siano soddisfatte le tre condizioni solite ($b \geq 0$, $m < n$, $r(A) = m$).
Sia B una matrice $m * m$ non singolare (matrice di base). Allora $x = B^{-1}b$ è soluzione di base. Se $x \geq 0$, abbiamo una soluzione ammissibile di base.
Possiamo scrivere: $x_B = B^{-1}b - B^{-1}Fx_F$
Da cui: $cx = c_B B^{-1}b + (c_F - c_B B^{-1}F)x_F = COST + \hat{c}x$
Dove $\hat{c} = [\hat{c}_B \ \hat{c}_F] = [0 \ c_F - c_B B^{-1}F]$ sono i cosiddetti costi ridotti. Notare che i coefficienti di costo ridotto corrispondenti alle variabili di base sono tutti *nulli*.
Teorema. Sia dato un problema di PL in forma standard, che soddisfa le 3 condizioni. Se per una soluzione ammissibile di base si ha $\hat{c} \leq 0$ (ovvero tutti i corrispondenti costi ridotti sono minori o uguali di zero), allora è una soluzione ottima.

Il metodo del simplesso attraverso un esempio.
$\max x_1 + x_2$
$6x_1 + 4x_2 \leq 24$
$3x_1 - 2x_2 \leq 6$
$x_1, x_2 \geq 0$
Convertiamo in forma standard aggiungendo variabili di slack x_3 e x_4:
$\max x_1 + x_2$
$6x_1 + 4x_2 + x_3 \qquad = 24$
$3x_1 - 2x_2 \qquad + x_4 = 6$
$x_1, x_2, x_3, x_4 \geq 0$
Abbiamo matrice A, con 6 sottomatrici $2 * 2$ non singolari.
Otteniamo 6 soluzioni di base, di cui 4 ammissibili e 2 no.
Partiamo dalla base x_3, x_4.

Otteniamo soluzione ammissibile di base che è $(0, 0, 24, 6)$, corrispondente al punto A nel disegno fatto alla lavagna, ovvero all'origine degli assi nel problema originario, in due variabili.
I costi ridotti sono uguali a quelli normali (ciò è dovuto al fatto che la matrice di base è l'identità e al fatto che sto usando come variabili di base tutte e sole le variabili di slack): $\hat{c} = (1, 1, 0, 0)$.
Se fosse $\hat{c} \leq 0$, STOP: abbiamo una soluzione ottima. Nel nostro esempio, non è $\hat{c} \leq 0$. Allora proviamo a cambiare base. Mettiamo in base una delle variabili corrispondente a un coefficiente di costo ridotto strettamente positivo.
Possiamo mettere in base x_1 oppure x_2. Scegliamo di mettere in base x_1.
Naturalmente, dovremo togliere una delle variabili che erano in base. Riscriviamo il sistema dei vincoli (mi riferisco ai vincoli $Ax = b$) tenendo uguale a zero le variabili che erano fuori base e che non facciamo entrare in base a questo passo (nel nostro esempio si tratta di x_2):
$+x_3 \qquad = 24 - 6x_1$
$+x_4 = 6 - 3x_1$
Se prendiamo $x_1 = 24/6$ si annulla x_3. Ma x_4 sarà negativa e quindi non avremo una soluzione ammissibile di base. (Nel disegno fatto alla lavagna troviamo il punto H)
Se prendiamo $x_1 = 6/3$ si annulla x_4. E la x_3 resta non negativa. Quindi rimaniamo con una soluzione *ammissibile* di base.
In pratica, in generale, dobbiamo guardare i coefficienti *positivi* che compaiono nella colonna della matrice sotto la variabile che vogliamo far entrare in base. Tra le righe corrispondenti a questi coefficienti cerchiamo la riga per cui il quoziente fra il termine noto e il relativo coefficiente nella matrice è minimo. La variabile da far uscire dalla base è quella corrispondente a questa riga (se vi fossero più righe che rendono minimo il rapporto, scegliamo una tra queste, a nostro piacimento).
Se per caso tutti i coefficienti della matrice nella colonna della variabile che abbiamo deciso di far entrare in base fossero minori o uguali a zero? Vorrebbe dire che la funzione obiettivo non è superiormente limitata. Quindi, STOP: il problema dato non ha soluzioni ottime.
A questo punto si effettua la operazione di pivoting per far sì che la matrice corrispondente alle nuove variabili di base sia la matrice identità.
Da qui si riparte per il passo successivo. E si va avanti fino a che non si arriva allo STOP (che può esserci perché abbiamo trovato una soluzione ottima oppure perché sappiamo che la funzione obiettivo non è superiormente limitata.

Resta un dubbio da sciogliere: siamo sicuri che questo algoritmo termini dopo un numero finito di passi? La risposta è no. Saremmo sicuri che l'algoritmo termina dopo un numero finito di passi se sapessimo che non si ripassa mai da un vertice già visitato. Questa eventualità non la possiamo escludere. Può verificarsi un caso di cycling, che può occorrere in corrispondenza di soluzioni di base degeneri (ovvero con qualche coordinata nulla; questo avviene ad esmepio se si trova che (almeno) due variabili rendono minimo il rapporto che ci serviva per trovare la variabile da far uscire dalla base). Il problema del cycling è più di carattere teorico che pratico. Comunque, vi è una regola (di Bland) che può essere usata se si vuole eliminare questo rischio.

Vediamo il tableau:

```
      x_1  x_2  x_3  x_4

       1    1    0    0

x_3    6    4    1    0    24
x_4    3   -2    0    1     6
```

Faccio entrare in base x_1. Calcolo $24/6 = 4$ e $6/3 = 2$ per scoprire chi fare uscire. Esce x_4 perché è sulla sua riga che trovo il quoziente più piccolo.
Quindi entra x_1 ed esce x_4.
Operazione di pivoting per fare comparire una colonna (sotto ad x_1) di tutti zeri (compresi i coefficienti della funzione obiettivo) tranne che in corrispondenza del pivot dove dovrà comparire 1.
Divido per 3 la riga corrispondente a x_4 (notare che le operazioni le faccio anche sui termini noti) per far comparire 1 in corrispondenza del pivot:

```
      x_1  x_2  x_3  x_4

       1    1    0    0

       6    4    1    0    24
       1  -2/3   0   1/3   6/3
```

Moltiplico per -2 la riga del pivot e la sommo alla riga x_3 in modo che compaia uno zero:

```
      x_1  x_2  x_3  x_4

       1    1    0    0
```

```
 0    8    1   -2    24
 1  -2/3   0   1/3   6/3
```

Moltiplico per $-1/3$ la riga del pivot e la sommo alla riga dei coefficienti c in modo che compaia uno zero:

```
x_1  x_2  x_3  x_4

 0   2/3   0   -1/3

 0    8    1   -2    12
 1  -2/3   0   1/3   6/3
```

Ho ottenuto un nuovo tableau in forma canonica, questa volta per la base x_1, x_3:

```
      x_1  x_2  x_3  x_4

       0   2/3   0   -1/3

x_3    0    8    1   -2    12
x_1    1  -2/3   0   1/3   6/3
```

Notare il modo in cui ho contrassegnato le righe!
Ho ancora un costo ridotto positivo, quindi faccio entrare in base la variabile corrispondente x_2). Visto che nella colonna di x_2 c'è solo un elemento strettamente positivo, sarà la variabile di base corrispondente a tale riga (e quindi x_3) ad uscire di base.
Avendo individuato il pivot, ricalcolo il nuovo tableau. Etc.
DOCUMENTI:
.
ESERCIZI:
.

Ve 13 ottobre 2006, 2h-20h:
Ripreso l'esempio di ieri e completati i conti sul tableau fino a trovare una soluzione ottima.
Un esempio con funzione obiettivo non superiormente illimitata.
Un esempio di soluzione di base degenere. Un passo dell'algoritmo del simplesso non fa spostare dal vertice in cui si era.

Metodo a due fasi per determinare una soluzione ammissibile di base.
Il problema ausiliario di PL ha come funzione obiettivo $\sum_{i=1}^{m} -y_i$. Per questo problema l'insieme dei vincoli, il poliedro P, è sempre non vuoto. Inoltre, la funzione obiettivo è superiormente limitata da 0 e quindi esiste sempre soluzione ottima. Se il valore della funzione obiettivo è 0, allora il problema di partenza aveva una soluzione ammissibile di base (e il metodo del simplesso, applicato al problema ausiliario, ce la fa trovare). Se invece il valore della funzione obiettivo è strettamente minore di zero, allora il problema dato non ha soluzione ammissibile di base.
DOCUMENTI:
Degenerazione:
pagine da Dantzig (fig. uno e due e breve testo)
pagine da Tadei e Della Croce (fig. uno due e tre).
ESERCIZI:
.

Gi 19 ottobre 2006, 3h-23h:
Un esempio di applicazione del metodo a due fasi. Determinazione dei costi ridotti per il problema ausiliario. La soluzione del problema ausiliario non ha in base le variabili ausiliarie e quindi ci dà direttamente una soluzione ammissibile di base del problema originario. Calcolo dei coefficienti di costo ridotti per il problema originario.

Dualità: duale di un problema di PL in forma canonica.
Dato:
$\max cx$
t.c.
$Ax \leq b$
$x \geq 0$
Il duale è:
$\min yb$
t.c.
$yA \geq c$
$y \geq 0$
Il duale del duale è il problema di partenza (detto primale).
Duale di un problema in forma standard.
Teorema debole della dualità ($cx \leq yb$ per x, y soluzioni ammissibili, rispettivamente per problema primale e duale) e dimostrazione.
Conseguenze di questo teorema.

Teorema forte della dualità. Se (P) ha soluzione ottima, allora:
- anche (D) ha soluzione ottima e max(P) = min(D)
- se $x^\star$ è soluzione ottima per (P), allora esiste $y^\star$ ottima per (D) e $cx^\star = y^\star b$.
Tabella riassuntiva di alcune conseguenze dei due teoremi di dualità

P \ D	vuoto	sup. illimit.	sol. ott.
vuoto	SI	SI	NO
sup. illimit.	SI	NO	NO
sol. ott.	NO	NO	SI

Un esempio di problemi duali. Disegno delle zone ammissibili.
Esempio di interpretazione del duale: il duale del problema della dieta.

Condizioni di ortogonalità (o complementary slackness).
Dati problema primale e duale (in forma canonica), siano $x^\star$ e $y^\star$ soluzioni di (P) e (D).
Abbiamo:
$Ax^\star \leq b$ e $x^\star \geq 0$ ammissibilità per il problema primale
$y^\star A \geq c$ e $y^\star \geq 0$ ammissibilità per il problema duale
$cx^\star = y^\star b$ teorema di dualità forte
Dalla dimostrazione del teorema di dualità debole abbiamo:
$cx^\star \leq y^\star Ax^\star \leq y^\star b$
Da tutto ciò deduciamo le condizioni di ortogonalità:
$y^\star(b - Ax^\star) = 0$ e $(y^\star A - c)x^\star = 0$
E quindi, ognuno degli addendi deve essere nullo. Quindi, se ad esempio fosse $(y^\star)_i > 0$, avremmo $(b - Ax^\star)_i = 0$.
Verifica di queste condizioni in un esempio.
DOCUMENTI:
.
ESERCIZI:
.

Ve 20 ottobre 2006, 2h-25h:
Tabellina di corrispondenze fra primale e duale (vedi Tadei e Della Croce, oppure Hiller e Lieberman).
Le condizioni:
$Ax^\star \leq b$ e $x^\star \geq 0$
$y^\star A \geq c$ e $y^\star \geq 0$
$cx^\star = y^\star b$
Sono anche sufficienti per garantire la ottimalità di $x^\star$ e $y^\star$.

Inoltre, se $\hat{x}$ e $\hat{y}$ soddisfano le condizioni di ortogonalità e sono anche soluzioni ammissibili (di primale e duale, rispettivamente), allora sono soluzioni ottimali.

Si può fare una analisi di *sensitività* della soluzione della equazione $\nabla f(x) = 0$ rispetto a parametri. Se abbiamo funzioni di classe $\mathcal{C}^2$ possiamo scrivere una condizione (vedi Bertsekas). Nel caso semplice, di una funzione di una variabile, dipendente da un parametro a, ovvero data $f(x, a)$, abbiamo: $\sigma'(a) = -\frac{\frac{\partial^2 f}{\partial a \partial x}(\sigma(a),a)}{\frac{\partial^2 f}{\partial x^2}(\sigma(a),a)}$.
Ho indicato con $\sigma'(a)$ la derivata della funzione $a \mapsto \sigma(a)$ che ad ogni valore del parametro a associa il valore, supposto unico, del punto critico
Si può fare una verifica con $f(x, a) = x^2 - 6ax$. Il punto che annulla il gradiente (qui abbiamo un punto di minimo) è $x = 3a$. Ovvero, $\sigma(a) = 3a$ E quindi $\sigma'(a) = 3$

Per la sensitività in PL, osserviamo che il teorema di dualità forte ci fornisce anche una informazione ulteriore. Se abbiamo una soluzione ottimale $x^\star = [(x^\star_B), (0)]$ relativa ad una base B, allora $y^\star = c_B B^{-1}$ è una soluzione ottimale del problema duale (questa è la strada usata da Tadei e Della Croce per dimostrare il teorema di dualità forte).
Da questa relazione e dalle condizioni di ortogonalità, scritte per un problema primale in forma standard, si può stimare la variazione della soluzione ottima e del valore ottimo al variare dei dati del problema di PL.
In particolare, se passiamo da b a $b + \Delta b$, la variazione del valore ottimo è:
$\Delta z = c_B B^{-1}(b + \Delta b) - c_B B^{-1} b = c_B B^{-1} \Delta b = y^\star \Delta b = \sum_{i=1}^m y_i^\star \Delta b_i$

Visti questi risultati nell'esempio solito (tratto da Fischetti).
DOCUMENTI:
.
ESERCIZI:
.

NON FATTO

Sensitività e interpretazione dei moltiplicatori di Lagrange. Per i dettagli delle ipotesi ed una formulazione corretta del risultato, vedi Bertsekas. Qui ci si limita a mettere l'accento sulla formula. Se consideriamo il problema di minimizzare $f(x)$ con il vincolo (vettoriale: $u \in \mathbb{R}^m$) $h(x) = u$, abbiamo che $\nabla p(u) = -\lambda(u)$. Dove $p(u) = f(x(u))$, con $x(u)$ il punto di minimo

del problema vincolato, in corrispondenza del parametro u. In particolare: $\nabla p(0) = -\lambda^\star$
C'è il significato, importante, dei moltiplicatori come prezzi ombra. Rapporto fra guadagno e la variazione di quantità disponibile di una risorsa vale λ.
Sensitività e interpretazione dei moltiplicatori di Kuhn-Tucker. Anche qui per i dettagli vedi Bertsekas. Se consideriamo il problema di minimizzare $f(x)$ con i vincoli $h(x) = u$ e $g(x) \leq v$, , abbiamo che $\nabla_u p(u, v) = -\lambda(u, v)$ e $\nabla_v p(u, v) = -\mu(u, v)$. Qui $p(u, v) = f(x(u, v))$, con $x(u)$ il punto di minimo del problema vincolato, in corrispondenza del parametro (u, v). In particolare: $\nabla_u p(0, 0) = -\lambda^\star$ e $\nabla_v p(0, 0) = -\mu^\star$
Notare, a proposito di prezzi ombra che i moltiplicatori dei vincoli di disuguaglianza non attivi valgono 0. Infatti, all'ottimo sto usando quella risorsa al di sotto della disponibilità che ne ho e quindi non sono interessato a pagare per averne di più.

A livello di intuizione, ho $g(x) \leq v$. Quindi se v aumenta il vincolo si allarga. Allora il min diminuisce. Cioè $p(v)$ diminuisce. Quindi è ragionevole pensare che:
$\nabla_v p(u, v) \leq 0$ se e solo se $-\mu_j^\star \leq 0$ (se e solo se $\mu_j^\star \geq 0$)

BIBLIOGRAFIA:
Una breve bibliografia.

Ritorna alla home page di Patrone