Corso:

Ricerca Operativa I (co-docenza)

LS Ingegneria dei Trasporti e della Logistica, Università di Genova, 2007/08

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

E' una versione provvisoria (e penso resterà tale per sempre).
Le ultime correzioni fatte sono evidenziate in rosso.
Per la precisione, le ultime con la parole **NEW ed in "bold"**. Le "penultime" solo in rosso.
Consultare la data/ora dell'ultimo aggiornamento (vedi in fondo).

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

Lu 24 settembre 2007, 1h-1h:
Introduzione al corso.
DOCUMENTI:
.
ESERCIZI:
.

Gi 11 ottobre 2007, 3h-4h:
Massimi e minimi, definizioni.
Parliamo di max (per min basta cambiare segno)
max 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 max globale e di (valore) massimo globale (dedotta da quella di punto di max).
Si può anche seguire strada opposta, partendo da max di $f(A)$.
Punto di max stretto.
Non ha senso parlare di max locale. Dobbiamo poter parlare di intorni. Prendiamo $A \subseteq \mathbb{R}^n$
$f : A \subseteq \mathbb{R}^n \to \mathbb{R}$.
Def di max locale: c'è intorno I (sferico, se vogliamo) di $\bar{x}$ t.c. $\bar{x}$ sia punto di max (globale) per $f|(I \cap A)$
Caso particolare $A \subseteq \mathbb{R}$. C'è $\delta > 0$ tale che, posto $I =]\bar{x} - \delta, \bar{x} + \delta[$, il punto $\bar{x}$ è punto di max globale per f ristretta a $I \cap A$.

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, per funzioni di una variabile. E' data $f : A \to \mathbb{R}$, con $A \subseteq \mathbb{R}$. E $\bar{x}$ è p.to di max 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 ≤ 0 se $\bar{x} < x$) .
3. usare teorema di permanenza del segno
Si ha allora $f'(\bar{x}) \leq 0$, "arrivando da destra". Se si arriva "da sinistra", $f'(\bar{x}) \geq 0$. Allora $f'(\bar{x}) = 0$. Ma attenzione, stiamo dimenticando cose importantissime!!!

Vediamo i dettagli.
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. Se lo è solo da una parte, possiamo comunque scrivere la condizione "unilaterale").
Sono comunque interessanti anche le condizioni "unilaterali" che, nel caso di un intervallo, si hanno negli estremi (se $A = [a, b]$, ed a è massimo, $f'(a) \leq 0$; se b è massimo, $f'(b) \geq 0$.
DOCUMENTI:
.
ESERCIZI:
.

Ve 12 ottobre 2007, 2h-6h:
Generalizzazione ad $\mathbb{R}^n$: sia $\bar{x}$ un p.to di max locale per $f : A \to \mathbb{R}$, con $A \subseteq \mathbb{R}^n$. Allora, dato un versore $v \in \mathbb{R}^n$ (si potrebbe anche prendere un vettore, basta che $v \neq 0$), $\lim_{t \to 0^+} \dfrac{f(x+tv) - f(x)}{t} \leq 0$ se ha senso.
Ovviamente, $f(x+tv)$ deve essere definita per "$t > 0$ sufficientemente piccoli". Vale a dire, abbiamo bisogno che il punto $\bar{x}$ sia di accumulazione per $A \cap S_{\bar{x}, v}$, dove $S_{\bar{x}, v}$ indica la semiretta uscente da $\bar{x}$, individuata dal versore v.
Quindi: $\dfrac{\partial f}{\partial v^+}(\bar{x}) \leq 0$ se ha senso. [Con $\dfrac{\partial f}{\partial v^+}(x)$ indichiamo il limite scritto sopra: $\lim_{t \to 0^+} \dfrac{f(x+tv) - f(x)}{t}$].

In particolare, se f è parzialmente derivabile ed $\bar{x}$ è interno ad A, per ogni $i \in \{1, \ldots, n\}$ si ha $\dfrac{\partial f}{\partial x_i}(\bar{x}) = 0$.

Questo fatto deriva facilmente dalla osservazione che (in due variabili, con n variabili i conti sono simili):
--- $\dfrac{\partial f}{\partial x^+}(\bar{x}) = \dfrac{\partial f}{\partial v^+}(\bar{x}) \leq 0$, prendendo come v il versore $(1, 0)$.
--- mentre $\dfrac{\partial f}{\partial x^-}(\bar{x}) = -\dfrac{\partial f}{\partial w^+}(\bar{x})$, dove w è il versore $(-1, 0)$. Essendo $\dfrac{\partial f}{\partial w^+}(\bar{x}) \leq 0$, otteniamo che $\dfrac{\partial f}{\partial x^-}(\bar{x}) \geq 0$.

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 concave!
Data $f \in \mathcal{C}^1(A)$, con A aperto e convesso di $\mathbb{R}^n$, se f è concava allora, per $\bar{x} \in A$:
$\nabla f(\bar{x}) = 0$ se e solo se $\bar{x}$ è punto di max assoluto per f.
Lo si dimostra usando la caratterizzazione differenziale della concavità: $f \in \mathcal{C}^1(A)$, con A aperto, è concava se e solo se per ogni $\bar{x}, x \in A$ si ha $f(x) \leq f(\bar{x}) + \nabla f(\bar{x}) \cdot (x - \bar{x})$.

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 max. Cerco di provare se queste condizioni

possano essere necessarie perché la f abbia max (o sufficienti, magari "rafforzandole un poco": ad esempio, chiedendo che la funzione approssimante abbia un punto di max *stretto*

Supponiamo che sia $f \in \mathcal{C}^2(A)$

Noi approssimeremo f con il polinomio di Taylor (del secondo ordine):

$$f(x) + \sum_{i=1,\ldots,n} \frac{\partial f}{\partial x_i}(\bar{x})(x_i - \bar{x}_i) + \sum_{i,j=1,\ldots,n} \frac{\partial^2 f}{\partial x_i x_j}(\bar{x})(x_i - \bar{x}_i)(x_j - \bar{x}_j).$$

Per provare quanto affermeremo, occorre naturalmente provare che il "resto" della formula di Taylor può essere davvero trascurato, ai nostri fini. Questo si ottiene esprimendo tale esito nella forma di Peano e sfruttando il fatto che va a zero più rapidamente del quadrato della distanza di x da $\bar{x}$.

La parte su cui concentrare l'attenzione è la forma quadratica: $\sum_{i,j=1,\ldots,n} \frac{\partial^2 f}{\partial x_i x_j}(\bar{x})(x_i - \bar{x}_i)(x_j - \bar{x}_j)$. Essa ha max in $x = \bar{x}$

Per comodità chiamiamo $\Phi(h) = \sum_{i,j=1,\ldots,n} \frac{\partial^2 f}{\partial x_i x_j}(\bar{x}) h_i h_j$ e vediamo se Φ ha max in 0 .

Chiaramente Φ ha max in 0 se e solo se $\Phi(h) \leq 0$ per ogni $h \in \mathbb{R}^n$.

se $\Phi(h) \leq 0$ per ogni $h \in \mathbb{R}^n$, si dice che Φ è semidefinita negativa..

Come verificare se Φ è semidefinita negativa? Con la matrice hessiana: $(H_{ij})_{i,j=1}^n = \frac{\partial^2 f}{\partial x_i x_j}(\bar{x})$. E' una matrice simmetrica ($f \in \mathcal{C}^2(A)$ e quindi vale il teorema di Schwartz). Quindi diagonalizzabile. Si dimostra che Φ sarà semidefinita negativa se e solo se tutti i suoi autovalori saranno minori o uguali a zero.

Invece, la condizione $\Phi(h) < 0$ per ogni $h \neq 0$, cioè se Φ abbia in 0 un punto di max stretto, è una condizione sufficiente perché $\bar{x}$ sia un punto di max locale (ammesso che sia anche $\nabla f(\bar{x}) = 0$).

Se $\Phi(h) > 0$ per ogni $h \neq 0$, diciamo che Φ è definita negativa.

E Φ è definita negativa se e solo se ogni suo autovalore è strettamente negativo

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

Problemi di max 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.

DOCUMENTI:

.

ESERCIZI:

.

Gi 18 ottobre 2007, 3h-9h:

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.

Faremo poi sempre l'ipotesi che A sia un aperto di $\mathbb{R}^n$.

Il teorema di Weierstrass garantisce esistenza di p.to di massimo (globale). Ipotesi: $f : A \to \mathbb{R}$, $A \subseteq \mathbb{R}^n$ chiuso e limitato, f continua. In queste ipotesi siamo certi che f abbia massimo (e quindi almeno un punto di max) e, anche, se ci servisse, che abbia minimo.

C'è anche teorema di Weierstrass generalizzato: se A non è limitato (pur essendo chiuso), basta che $\lim_{\|x\| \to \infty} f(x) = -\infty$, per essere certi della esistenza di un punto di max globale. (Se ci interessassero i min, dovremmo sostituire la condizione: $\lim_{\|x\| \to \infty} f(x) = -\infty$ con $\lim_{\|x\| \to \infty} f(x) = +\infty$)

Massimi e minimi vincolati (al solito, parleremo solo di max).
Abbiamo già visto il "setting". Abbiamo $A \subseteq \mathbb{R}^n$, e sono date f, h_i, con $i = 0, \ \ldots, m$ e g_j, $j = 0, \ \ldots, r$. Tutte definite su A e a valori in $\mathbb{R}$
Cerchiamo p.ti di max (locale, globale) per $f|V$ dove V è l'insieme dei punti che soddisfano il vincolo (o, brevemente, il vincolo). Cioè è l'insieme dei punti $x \in A$ 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 massimo vincolato, abbiamo la condizione detta "dei moltiplicatori di Lagrange".
Occorre assumere anche che i vettori $\nabla h_i(x^\star)$ siano *linearmente indipendenti*.
Se questa condizione non è soddisfatta, non abbiamo garanzia che la tesi del teorema dei moltiplicatori di Lagrange valga. Ciò significa, in parole povere, che punti (in E, ovvero che stiano sul vincolo!) nei quali i vettori $\nabla h_i(x^\star)$ non sono linearmente indipendenti andranno sempre messi nella lista dei sospettabili di essere punti di max locale.

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; l'unicità è una banale conseguenza della ipotesi di indipendenza lineare dei $\nabla h_i(x^\star)$) 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 $x^\star$ soddisfa anche la condizione $h(x^\star) = 0$ (cioè: $h_i(x^\star) = 0$ 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:.

$$\begin{cases} \nabla_x L(x^\star, \lambda^\star) = 0 \\ \nabla_\lambda L(x^\star, \lambda^\star) = 0 \end{cases}$$

Mi riferirò a questo sistema, se necessario, come al sistema (L).

ESEMPIO:
$f(x, y) = x$ con vincolo $x + y = 1$. Le condizioni del teorema sono soddisfatte (in particolare, l'unico vettore $\nabla h(x, y)$ è costante ed uguale a $(1, 1)$. Quindi non si annulla mai e pertanto neanche sul vincolo E. Pertanto l'insieme dei vettori $\nabla h(x, y)$ è sempre un insieme linearmente indipendente (per un insieme che contiene un solo vettore ciò equivale a verificare che questo vettore non sia il vettore nullo).
Non troviamo nessun punto che risolva il sistema (L).
Il sistema (L), infatti, è:

$$\begin{cases} 1-\lambda=0 \\ -\lambda=0 \\ x-\lambda(x+y-1)=0 \end{cases}$$

E, chiaramente, le prime due equazioni sono contraddittorie fra loro.
D'altronde, è evidente "a occhio" che il problema dato non presenta alcun punto di max locale.

. ESEMPIO:
$f(x, y) = x$ con vincolo $x^2 + y^2 = 1$. Le condizioni del teorema sono soddisfatte (in particolare, l'unico vettore $\nabla h(x, y)$ vale $(2x, 2y)$ e quindi non si annulla sul vincolo).
Il sistema (L) è:

$$\begin{cases} 1-2\lambda x=0 \\ -2\lambda y=0 \\ x^2+y^2=1 \end{cases}$$

Questo sistema ha due soluzioni, le cui coordinate x ed y sono, rispettivamente: $(1, 0)$ e $(-1, 0)$.
Usando il teorema di Weierstrass si vede che uno, il punto $(1, 0)$, è punto di max globale. L'altro è punto di min globale.

Senza la condizione di lineare indipendenza, il teorema è falso. Prendere: $f(x, y, z) = x + y + z$ (ad esempio, ma c'è ampia scelta...) e i vincoli siano $h_1(x, , y, z) = x^2 + y^2 + (z-1)^2 - 1 = 0$ e $h_1(x, , y, z) = x^2 + y^2 + (z-2)^2 - 4 = 0$. Il vincolo descrive un singleton. Ma il sistema (L) non ha soluzione:

$$\begin{cases} 1+2\lambda x+2\mu x=0 \\ 1+2\lambda y+2\mu y=0 \\ 1+2\lambda(z-1)+2\mu(z-2)=0 \\ x^2+y^2+(z-1)^2=1 \\ x^2+y^2+(z-2)^2=4 \end{cases}$$

Infatti dalle due ultime equazioni ricaviamo che $(x, y, z) = (0, 0, 0)$. Sostituendo nella prima equazione troviamo $1 = 0$.
Tutto normale: non essendo soddisfatte le condizioni del teorema dei moltiplicatori di Lagrange non abbiamo nessuna garanzia che il sistema (L) abbia soluzione (anche se la potrebbe aver avuta, in linea generale).
Se abbiamo anche vincoli di disuguaglianza, ci sono le condizioni di Kuhn-Tucker.
Stavolta il vincolo è dato da $h_i(x) = 0$ e $g_j(x) \le 0$
Serve l'idea di vincolo *attivo*. Un vincolo di disuguaglianza (il vincolo j) è attivo in un punto $x^\star$ se $g_j(x^\star) = 0$.
Indichiamo con $A(x^\star)$ l'insieme degli *indici* [nota bene: $A(x^\star) \subseteq \{1, \ldots, r\}$] che identificano i vincoli attivi nel punto $x^\star$
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. (introducendo la lagrangiana:
$L(x, \lambda, \mu) = f(x) - \sum_{i=1}^{m} \lambda_i h_i(x) - \sum_{j=1}^{r} \mu_j g_j(x)$):

$$\begin{cases} \nabla_x L(x^\star, \lambda^\star, \mu^\star) = 0 \\ \nabla_\lambda L(x^\star, \lambda^\star, \mu^\star) = 0 \\ \nabla_\mu L(x^\star, \lambda^\star, \mu^\star) \geq 0 \\ \mu^\star \geq 0 \\ \mu^\star \cdot \nabla_\mu L(x^\star, \lambda^\star, \mu^\star) = 0 \end{cases}$$

L'ultima condizione vuol dire: $\mu_j^\star \cdot g_j(x^\star) = 0$ per ogni j

La seconda e terza equazione nel sistema non sono altro che le condizioni dei vincoli. Possiamo allora riscrivere il sistema così

$$\begin{cases} \nabla_x L(x^\star, \lambda^\star, \mu^\star) = 0 \\ h(x^\star) = 0 \\ g(x^\star) \leq 0 \\ \mu^\star \geq 0 \\ \mu^\star \cdot g(x^\star) = 0 \end{cases}$$

DOCUMENTI:

.

ESERCIZI:

.

Ve 19 ottobre 2006, 2h-11h:

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.

ESEMPIO:

$f(x, y) = x$ con vincolo $x^2 + y^2 \leq 1$. Vediamo cosa succede per $(x, y) = (1, 0)$, che è chiaramente punto di max assoluto per f col vincolo dato. Il vincolo è attivo in $(1, 0)$. Le condizioni del teorema sono soddisfatte (in particolare, l'unico vettore $\nabla g(x, y)$ vale $(2x, 2y)$ e quindi non si annulla in $(1, 0)$).

Il sistema (L) è:

$$\begin{cases} 1 - 2\mu x = 0 \\ -2\mu y = 0 \\ x^2 + y^2 = 1 \\ \mu \geq 0 \end{cases}$$

Ovviamente $\left(1, 0, \frac{1}{2}\right)$ è soluzione del sistema. Notare che il moltiplicatore vale $\frac{1}{2}$ (per K-T doveva essere maggiore o uguale a zero).

ESEMPIO:
$f(x, y) = x$ con vincolo $x^2 + y^2 \leq 1$ e $x + y \leq 5$. Vediamo cosa succede per $(x, y) = (1, 0)$, che è chiaramente punto di max assoluto per f col vincolo dato. Il primo vincolo è attivo in $(1, 0)$, il secondo no. Le condizioni del teorema sono soddisfatte (in particolare, l'unico vettore $\nabla g_1(x, y)$ che ci interessa, in quanto corrispondente all'unico vincolo attivo, vale $(2x, 2y)$ e quindi non si annulla in $(1, 0)$).
Il sistema (L) è:

$$\begin{cases} 1 - 2\mu_1 x - \mu_2 = 0 \\ -2\mu y - \mu_2 = 0 \\ x^2 + y^2 = 1 \\ x + y < 5 \\ \mu_1 \geq 0 \\ \mu_2 = 0 \end{cases}$$

Ovviamente $\left(1, 0, \frac{1}{2}, 0\right)$ è soluzione del sistema. Notare che il primo moltiplicatore vale $\frac{1}{2}$ (per K-T doveva essere maggiore o uguale a zero), mentre il secondo vale 0.

ESEMPIO SVOLTO:
$\max\ x + 2y + 3z$ 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.

Osserviamo che sia la funzione f data ($f(x, y, z) = x + 2y + 3z$) che le funzioni g_1 e g_2 che descrivono i vicoli sono definite su $\mathbb{R}^2$ e sono anche di classe $C^1(\mathbb{R}^2)$.
Per quanto riguarda la lineare indipendenza dei vincoli, notiamo che:
$g_1(x, y, z) = x^2 + y^2 + z^2 - 4$ e $g_2(x, y, z) = 1 - z$ (per g_2 occorre fare attenzione che il vincolo era stato dato in forma di ≥ 0 invece che di ≤ 0).
$\nabla g_1(x, y, z) = (2x, 2y, 2z)$ e $\nabla g_2(x, y, z) = (0, 0, -1)$. Eccettuato il punto $(x, y, z) = (0, 0, 0)$ (che d'altronde non sta sul vincolo), negli altri punti e quindi sul vincolo i due vettori sono linearmente indipendenti.
Allora le condizioni di KT sono necessarie. Scriviamo il sistema che ci dà le condizioni di KT (la lagrangiana è:
$L(x, y, z, \mu_1, \mu_2) = f(x, y, z) - \mu_1 g_1(x, y, z) - \mu_2 g_2(x, y, z) = x + 2y + 3z + \mu_1(x^2 + y^2 + z^2 - 4) - \mu_2(1 - z)$
):

$$\begin{cases} \nabla_x L(x, y, z, \mu_1, \mu_2) = 0 \\ \dfrac{\partial L(x, y, z, \mu_1, \mu_2)}{\partial \mu_1} \leq 0 \\ \dfrac{\partial L(x, y, z, \mu_1, \mu_2)}{\partial \mu_2} \leq 0 \\ \mu_1 \geq 0 \\ \mu_2 \geq 0 \\ \mu_1 \dfrac{\partial L(x, y, z, \mu_1, \mu_2)}{\partial \mu_1} = 0 \\ \mu_2 \dfrac{\partial L(x, y, z, \mu_1, \mu_2)}{\partial \mu_2} = 0 \end{cases}$$

Chiamo $disuguaglianze\ (D)$ il seguente sistema:

$$\begin{cases} \dfrac{\partial L(x, y, z, \mu_1, \mu_2)}{\partial \mu_1} \leq 0 \\ \dfrac{\partial L(x, y, z, \mu_1, \mu_2)}{\partial \mu_2} \leq 0 \\ \mu_1 \geq 0 \\ \mu_2 \geq 0 \end{cases}$$

Usando la legge dell'annullamento del prodotto, possiamo dire che le soluzioni del sistema di KT sono date dalle soluzioni di almeno uno dei 4 sistemi seguenti, che indichiamo con (1), (2), (3) e (4).

Sistema (1):

$$\begin{cases} \nabla_x L(x, y, z, \mu_1, \mu_2) = 0 \\ \mu_1 = 0 \\ \mu_2 = 0 \\ disuguaglianze\ (D) \end{cases}$$

Sistema (2):

$$\begin{cases} \nabla_x L(x, y, z, \mu_1, \mu_2) = 0 \\ \mu_1 = 0 \\ \dfrac{\partial L(x, y, z, \mu_1, \mu_2)}{\partial \mu_2} = 0 \\ disuguaglianze\ (D) \end{cases}$$

Sistema (3):

$$\begin{cases} \nabla_x L(x,\ y,\ z,\ \mu_1,\ \mu_2) = 0 \\ \dfrac{\partial L(x,\ y,\ z,\ \mu_1,\ \mu_2)}{\partial \mu_1} = 0 \\ \mu_2 = 0 \\ disuguaglianze\ (D) \end{cases}$$

Sistema (4):

$$\begin{cases} \nabla_x L(x,\ y,\ z,\ \mu_1,\ \mu_2) = 0 \\ \dfrac{\partial L(x,\ y,\ z,\ \mu_1,\ \mu_2)}{\partial \mu_1} = 0 \\ \dfrac{\partial L(x,\ y,\ z,\ \mu_1,\ \mu_2)}{\partial \mu_2} = 0 \\ disuguaglianze\ (D) \end{cases}$$

Riscriviamoli in dettaglio nel nostro caso:

Sistema (1):

$$\begin{cases} 1 - 2\mu_1 x = 0 \\ 2 - 2\mu_1 y = 0 \\ 3 - 2\mu_1 z + \mu_2 = 0 \\ \mu_1 = 0 \\ \mu_2 = 0 \\ disuguaglianze\ (D) \end{cases} \quad \text{ovvero} \quad \begin{cases} 1 = 0 \\ 2 = 0 \\ 3 = 0 \\ \mu_1 = 0 \\ \mu_2 = 0 \\ disuguaglianze\ (D) \end{cases}$$

Ovviamente non ha soluzioni (nessun punto "interno al vincolo" soddisfa le CN).

Sistema (2):

$$\begin{cases} 1 - 2\mu_1 x = 0 \\ 2 - 2\mu_1 y = 0 \\ 3 - 2\mu_1 z + \mu_2 = 0 \\ \mu_1 = 0 \\ 1 - z = 0 \\ disuguaglianze\ (D) \end{cases} \quad \text{ovvero} \quad \begin{cases} 1 - \ = 0 \\ 2 - \ = 0 \\ 3 + \mu_2 = 0 \\ \mu_1 = 0 \\ 1 - z = 0 \\ disuguaglianze\ (D) \end{cases}$$

Anche per questo, ovviamente, nessuna soluzione (nessun punto sul piano di equazione $1 - z = 0$, e che appartenga al vincolo, soddisfa le CN).

Sistema (3):

$$\begin{cases} 1-2\mu_1 x = 0 \\ 2-2\mu_1 y = 0 \\ 3-2\mu_1 z + \mu_2 = 0 \\ x^2+y^2+z^2-4=0 \\ \mu_2 = 0 \\ \mathit{disuguaglianze}\ (D) \end{cases} \quad \text{ovvero} \quad \begin{cases} x = \dfrac{1}{2\mu_1} \\ y = \dfrac{2}{2\mu_1} \\ z = \dfrac{3}{2\mu_1} \\ \dfrac{1}{4\mu_1^2} + \dfrac{4}{4\mu_1^2} + \dfrac{9}{4\mu_1^2} = 4 \\ \mu_2 = 0 \\ \mathit{disuguaglianze}\ (D) \end{cases} \quad \text{ovvero} \quad \begin{cases} x = \dfrac{1}{2\mu_1} \\ y = \dfrac{2}{2\mu_1} \\ z = \dfrac{3}{2\mu_1} \\ \mu_1 = \pm \dfrac{\sqrt{7}}{8} \\ \mu_2 = 0 \\ \mathit{disuguaglianze}\ (D) \end{cases} \quad \text{ovvero}$$

$$\begin{cases} x = \dfrac{1}{2}\dfrac{\sqrt{8}}{7} \\ y = \dfrac{2}{2}\dfrac{\sqrt{8}}{7} \\ z = \dfrac{3}{2}\dfrac{\sqrt{8}}{7} \\ \mu_1 = \dfrac{\sqrt{7}}{8} \\ \mu_2 = 0 \\ \mathit{disuguaglianze}\ (D) \end{cases}$$

Nell'ultimo passaggio abbiamo usato le $\mathit{disuguaglianze}\ (D)$ per eliminare la soluzione corrispondente a $\mu_1 = -\frac{\sqrt{7}}{8}$. L'unica soluzione è quindi: $\left(\frac{1}{2}\frac{\sqrt{8}}{7}, \frac{\sqrt{8}}{7}, \frac{3}{2}\frac{\sqrt{8}}{7}, \frac{\sqrt{7}}{8}, 0\right)$.

Sistema (4):

$$\begin{cases} 1-2\mu_1 x = 0 \\ 2-2\mu_1 y = 0 \\ 3-2\mu_1 z + \mu_2 = 0 \\ x^2+y^2+z^2-4=0 \\ 1-z=0 \\ \mathit{disuguaglianze}\ (D) \end{cases} \quad \text{ovvero} \quad \begin{cases} x = \dfrac{1}{2\mu_1} \\ y = \dfrac{2}{2\mu_1} \\ 3-2\mu_1+\mu_2 = 0 \\ x^2+y^2 = 3 \\ z = 1 \\ \mathit{disuguaglianze}\ (D) \end{cases} \quad \text{ovvero} \quad \begin{cases} x = \dfrac{1}{2\mu_1} \\ y = \dfrac{2}{2\mu_1} \\ 3-2\mu_1+\mu_2 = 0 \\ \dfrac{1}{4\mu_1^2} + \dfrac{4}{4\mu_1^2} = 3 \\ z = 1 \\ \mathit{disuguaglianze}\ (D) \end{cases} \quad \text{ovvero}$$

$$\begin{cases} x = \dfrac{1}{2\mu_1} \\ y = \dfrac{2}{2\mu_1} \\ 3 - 2\mu_1 + \mu_2 = 0 \\ \mu_1 = \pm\dfrac{\sqrt{5}}{12} \\ z = 1 \\ disuguaglianze\ (D) \end{cases} \quad \text{ovvero} \quad \begin{cases} x = \dfrac{1}{2\mu_1} \\ y = \dfrac{2}{2\mu_1} \\ 3 - 2\mu_1 + \mu_2 = 0 \\ \mu_1 = \dfrac{\sqrt{5}}{12} \\ z = 1 \\ disuguaglianze\ (D) \end{cases} \quad \text{ovvero} \quad \begin{cases} x = \dfrac{1}{2\mu_1} \\ y = \dfrac{2}{2\mu_1} \\ \mu_2 = 2\dfrac{\sqrt{5}}{12} - 3 \\ \mu_1 = \dfrac{\sqrt{5}}{12} \\ z = 1 \\ disuguaglianze\ (D) \end{cases}$$

Si perviene al penultimo sistema grazie alle *disuguaglianze* (D) che escludono la soluzione corrispondente a $\mu_1 = -\dfrac{\sqrt{5}}{12}$. Nell'ultimo sistema è evidente che μ_2 non è ≥ 0, e quindi concludiamo che il sistema (4) non ha soluzioni.

Il sistema (4) mostra un fenomeno interessante. Rende evidente che quanto si fa non corrisponde alla ricerca di punti che soddisfino la CN di max locale ristretti ai vincoli: $x^2 + y^2 + z^2 - 4 = 0$ e $1 - z = 0$. In questo caso, vi sarebbe per forza un punto di max globale (Weierstrass). Mentre il sistema (4) non ci fornisce nessuna soluzione. L'apparente "mistero" è facilmente risolubile. Il sistema (4) cerca le soluzioni **del pb dato** che **si trovino sul vincolo** $x^2 + y^2 + z^2 - 4 = 0$ e $1 - z = 0$. **Non** cerca i punti di max di f **ristretta** al vincolo $x^2 + y^2 + z^2 - 4 = 0$ e $1 - z = 0$.
La differenza si vede benissimo nel seguente esempio:
$\max z$ con vincoli: $x^2 + y^2 - z \leq 0$ e $1 - z \leq 0$ (cioè, i punti dello spazio che stanno sia sopra al paraboloide di sezione circolare $z = x^2 + y^2$ che al piano di equazione $z = 1$).
Ovviamente tutti i punti della circonferenza (C) individuata (nello spazio) da $z = x^2 + y^2$ e $z = 1$ sono punti di massimo per f ristretta a tale vincolo (la f è costante e vale 1 sulla circonferenza (C)). Mentre, se scriviamo l'equivalente del sistema (4) in questo caso, vediamo che non ha soluzioni, come mostrato sotto. E questo è corretto, perché se ci si sposta verso l'alto, da un qualunque punto della circonferenza (C), il valore della funzione obiettivo aumenta e si resta dentro al vincolo dato. Quindi, il problema dato ovviamente non ha punti di max locale sulla circonferenza (C).
Sistema (4):

$$\begin{cases} 2\mu_1 x = 0 \\ 2\mu_1 y = 0 \\ 1 + \mu_1 + \mu_2 = 0 \\ x^2 + y^2 - z = 0 \\ 1 - z = 0 \\ disuguaglianze\ (D) \end{cases} \quad \text{ovvero} \quad \left[\begin{cases} \mu_1 \neq 0 \\ x = 0 \\ y = 0 \\ 1 + \mu_1 + \mu_2 = 0 \\ z = 0 \\ z = 1 \\ disuguaglianze\ (D) \end{cases} \quad vel \quad \begin{cases} \mu_1 = 0 \\ 0 = 0 \\ 0 = 0 \\ \mu_2 = -1 \\ 1 = x^2 + y^2 \\ z = 1 \\ disuguaglianze\ (D) \end{cases} \right]$$

Come si vede, nessuno dei due sistemi dentro la parentesi quadra ha soluzioni. In particolare, il secondo avrebbe come soluzione tutti i punti della circonferenza (C) se non ci fosse la condizione di non negatività sul moltiplicatore μ_2, che ci sta a ricordare che abbiamo "libertà di movimento" verso l'alto!

DOCUMENTI:
.
ESERCIZI:
.

Gi 25 ottobre 2007, non fatta (ore "cedute" a Di Febraro):

Ve 26 ottobre 2007, 2h-13h:
Interpretazione geometrica:
Distinzione fra Lagrange e K-T:
-- Lagrange: $\nabla_x f = \lambda_1 \nabla_x h_1 + \lambda_2 \nabla_x h_2$. Il gradiente di f è nello spazio vettoriale generato da $\nabla_x h_1$ e $\nabla_x h_2$. (O, anche, il gradiente di f è ortogonale allo spazio vettoriale che ci dà la approssimazione lineare [locale] dei vincoli).
---- esempio (con due vincoli, in $\mathbb{R}^3$): $f(x, y, z) = ax + by + cz$, $h_1(x, y, z) = x$, $h_2(x, y, z) = y$: pertanto il vincolo è $x = 0, \ y = 0$, e quindi è l'asse delle z. Quindi $\nabla f(x, y, z) = (a, b, c)$. Solo se $c = 0$, il gradiente di f sta nel piano xy, che è lo spazio vettoriale generato da $\nabla h_1 = (1, 0, 0)$ e da $\nabla h_2 = (0, 1, 0)$.

-- K-T: $\nabla_x f = \mu_1 \nabla_x g_1 + \mu_2 \nabla_x g_2$. Con $\mu_1, \ \mu_2 \geq 0$. Il gradiente di f è nel **cono** generato da $\nabla_x g_1$ e $\nabla_x g_2$.
---- esempio (con due vincoli, in $\mathbb{R}^3$): $f(x, y, z) = ax + by + cz$, $g_1(x, y, z) = x$, $g_2(x, y, z) = y$. Pertanto il vincolo è $x \leq 0, \ y \leq 0$, cioè "l'angolo diedro" individuato, nel piano xy, dal terzo quadrante. Quindi $\nabla f(x, y, z) = (a, b, c)$. Solo se $c = 0$, il gradiente di f sta nel **primo quadrante** del piano xy, che è il cono generato da $\nabla g_1 = (1, 0, 0)$ e da $\nabla g_2 = (0, 1, 0)$.

Algoritmi.
Criteri d'arresto (e loro problemi).
Metodo del gradiente.
Metodo di Newton.

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 2 novembre 2007, 4h-17h.
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 max (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.

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.
Ma abbiamo da massimizzare una funzione concava 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. 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 fra 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 quel vertice 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: ad x sostituisco $u - v$, con $u,\ v \geq 0$. Un po' la stessa idea la quale mi dice che $x = x^+ - x^-$. Anzi, in realtà se posso scrivere x come $u - v$, con $u,\ v \geq 0$, se è $u \geq v$, ho che $x^+ = u - v$ e $x^- = v - v = 0$; similmente se $u < v$ si ha $x^+ = 0$ e $x^- = v - u$. Come si può notare, $u,\ v$ non sono univocamente determinate da x)
$\max cx$
$Ax \leq b$
$x \geq 0$
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 $j \in \{1,\ \ldots,\ m\}$.
- $m \leq n$
- $r(A) = m$

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.
Variabili di slack.

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.
[Motzkin] Ogni poliedro P può essere decomposto così: $P = \mathrm{conv}(X) + \mathrm{cone}(Y)$, 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 $m \cdot m$ (cioè con m righe ed m colonne) B 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 $Bx_B = b - F\overline{x}_F$ ha soluzione unica comunque fissiamo $\overline{x}_F$.
Prima osservazione.
Se fissiamo $\overline{x}_F = 0$, 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).

DOCUMENTI:
Vedi Tadei-Della Croce per esempi di modellizzazione mediante problemi di progammazione lineare (per dettagli, vedi bibliografia).
ESERCIZI:
.

Gi 8 novembre 2007, 3h-20h:
Variabili di slack e di surplus. E ricordato come si trattano le variabili "libere" (cioè senza vincolo di segno).

Dato un problema in forma canonica, ad esso possiamo associare con una procedura sistematica un problema in forma standard, che gli è equivalente. Nel senso che per ogni soluzione ottima del problema in forma canonica possiamo individuare una corrispondente soluzione ottima del problema in forma standard e viceversa.
Vediamo. Abbiamo un problema in forma canonica:
$\max cx$
$Ax \leq b$
$x \geq 0$
Introduciamo le variabili $y \in \mathbb{R}^m$ e consideriamo il seguente problema, in forma standard.

$\max cx + 0y$
$Ax + Iy = b$
$x \geq 0$
$y \geq 0$
Proviamo che, se $\bar{x}$ è una soluzione ottima del problema in forma canonica, allora $(\bar{x},\ b - A\bar{x})$ è soluzione del problema in forma standard definito qui sopra.
E' immediato verificare che $(\bar{x},\ b - A\bar{x})$ soddisfa i vincoli. Resta da provare che è soluzione ottima.
Supponiamo per assurdo che vi sia un elemento $(\hat{x},\ \hat{y})$ che soddisfa anch'esso i vincoli: $A\hat{x} + I\hat{y} = b$
$\hat{x} \geq 0$
$\hat{y} \geq 0$
e t.c.:
$c\hat{x} + 0\hat{y} > c\bar{x} + 0(b - A\bar{x})$
Si vede subito che il vettore $\hat{x}$ soddisfa i vincoli del problema (in forma canonica) di partenza:
Da $A\hat{x} + I\hat{y} = b$, tenuto conto del fatto che $I\hat{y} = \hat{y}$ e $\hat{y} \geq 0$, abbiamo che $A\hat{x} \leq b$.
Inoltre, ovviamente è anche $\hat{x} \geq 0$.
Allora abbiamo un vettore $\hat{x}$ che soddisfa i vincoli del problema (in forma canonica) di partenza. Abbiamo anche supposto che $c\hat{x} + 0\hat{y} > c\bar{x} + 0(b - A\bar{x})$, ovvero che $c\hat{x} > c\bar{x}$. Ma allora $\hat{x}$ soddisfa i vincoli del problema in forma canonica e fa assumere alla funzione obiettivo un valore strettamente maggiore di quello che essa assume in $\hat{x}$. Contraddizione col fatto che $\hat{x}$ era una soluzione ottima del problema dato (in forma canonica).

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.

Occupiamoci ora di un problema di PL in forma standard, e supponiamo che siano soddisfatte le tre condizioni solite ($b \geq 0$, $m \leq n$, $r(A) = m$).
Sia B una matrice $m \cdot m$ 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: $cx = c_BB^{-1}b + (c_F - c_BB^{-1}F)x_F = COST + \hat{c}x$
Dove $\hat{c} = [\hat{c}_B\ \hat{c}_F] = [0\ c_F - c_BB^{-1}F]$ 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:
$\max x_1 + x_2$

$$\begin{cases} 6x_1 + 4x_2 + x_3 \quad\quad = 24 \\ 3x_1 - 2x_2 \quad\quad + x_4 = 6 \end{cases}$$

$x_1, x_2, x_3, x_4 \geq 0$

Abbiamo matrice A, con 6 sottomatrici $2 \cdot 2$ 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):

$$+x_3 \quad\quad = 24 - 6x_1$$
$$\quad\quad + x_4 = 6 - 3x_1$$

Se prendiamo $x_1 = \frac{24}{6}$ 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 $x_1 = \frac{6}{3}$ si annulla x_4. E la x_3 resta non negativa. Quindi rimaniamo con una soluzione *ammissibile* di base.
DOCUMENTI:
.
ESERCIZI:
Si consideri il problema $\max x + y$, con i vincoli $x \geq 2$ e $x \leq 3$. Trasformarlo in forma canonica e, in questa forma, trovare i vertici del poliedro dei vincoli. Idem con la forma standard.
Si consideri il problema $\max x + y$, con i vincoli $x + y \geq 2$ e $x + y \leq 3$. Trasformarlo in forma canonica e, in questa forma, trovare i vertici del poliedro dei vincoli. Idem con la forma standard. Osservare che, sia in forma canonica che in forma standard, assume un massimo in un vertice. A cosa corrisponde questo vertice, nel problema originario?

Ve 9 novembre 2007, 2h-22h:
Perché è naturale vedere c come "vettore riga", ovvero come matrice $1 \cdot n$ (è la matrice di rappresentazione di una applicazione lineare da $\mathbb{R}^n$ in $\mathbb{R}$).

Teorema fondamentale della PL Sia dato un problema di PL in forma standard. Indico con P il poliedro dei vincoli. Allora:
- se $P \neq \emptyset$, 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. Come esempi, vedansi gli esercizi già proposti.

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 esempio 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 $\frac{24}{6} = 4$ e $\frac{6}{3} = 2$ 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 $-\frac{1}{3}$ 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.

Un esempio con funzione obiettivo non superiormente limitata.
Un esempio di soluzione di base degenere. Un passo dell'algoritmo del simplesso non fa spostare dal vertice in cui si era.

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 15 novembre 2006, 3h-25h:
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.

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.

DOCUMENTI:
ESERCIZI:
.

Ve 16 novembre 2006, 2h-27h.
Dualità: duale di un problema di PL in forma canonica.
Dato:
$\max cx$
t.c. $\begin{cases} Ax \leq b \\ x \geq 0 \end{cases}$
Il duale è:
$\min yb$
t.c. $\begin{cases} yA \geq c \\ y \geq 0 \end{cases}$
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.

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.

Esempio di interpretazione del duale: il duale del problema di produzione.

DOCUMENTI:
.
ESERCIZI:
.

Ve 16 novembre 2006, 2h-30h:
Discussione e analisi del duale di un modello lineare di produzione, visto in dettaglio mediante un esempio.
Dimostazione del teorema di dualità forte (seguendo Tadei - Della Croce).
Tre condizioni equivalenti di ottimalità per primale e duale: la definizione, le condizioni di complementarità, l'uguaglianza dei valori.
Analisi di sensitività: breve introduzione di carattere generale (invito a comprendere per bene, calcoli inclusi, il caso più banale: come dipendono dai parametri a e b le soluzioni dlel'equazione $ax + b = 0$) e poi studio specifico della dipendenza dal "secondo membro" dei vincoli, b.
Per la sensitività in PL, osserviamo che il teorema di dualità forte ci fornisce anche una informazione ulteriore. Se abbiamo una soluzione ottimale $x^\star = \begin{bmatrix} x_B^\star \\ 0 \end{bmatrix}$ 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:

$$\Delta z = c_B B^{-1}(b + \Delta b) - c_B B^{-1} b = c_B B^{-1} \Delta b = y^\star \Delta b = \sum_{i=1}^{m} y_i^\star \Delta b_i$$

Significato di $y_i^\star$ come prezzi ombra, nel problema di produzione.

DOCUMENTI:
.
ESERCIZI:
.

BIBLIOGRAFIA:
Una breve bibliografia

Ultimo aggiornamento: 6 febbraio 2008

Ritorna alla home page di Patrone