Foto di Marco Bernardini

Matematici a dieta!

di Sonia De Cosmis *e* Renato De Leone

L'estate si avvicina e, con essa, anche la famigerata "prova costume": dopo l'esagerazione delle feste pasquali, con gli irrinunciabili piatti della nonna e quelle deliziose uova al cioccolato, è ora, ahimé, di mettersi a dieta!!!

Ciò che stiamo per affermare sembrerà strano, ma la matematica può fare al caso nostro e aiutarci a compilare una dieta equilibrata, senza rinunciare a ciò che più ci piace.

Il problema della dieta è uno dei primi problemi di ottimizzazione studiati negli Stati Uniti negli anni '30 e '40 del secolo scorso, per far fronte alle richieste nutritive del personale dell'esercito, minimizzando contemporaneamente il costo degli alimenti.
Uno dei primi ricercatori a occuparsi di questo problema fu George Stigler, che lo studiò e lo formulò nel 1939, identificando una soluzione con un metodo euristico. In seguito, nell'autunno del 1947, Jack Laderman risolse il problema con 9 equazioni e 77 variabili (vincolate ad assumere valori non negativi), impiegando circa 120 giorni/uomo per eseguire i calcoli manualmente, e scoprendo che la soluzione trovata precedentemente da Stigler era sbagliata per soli 24 centesimi!

Problemi come questo sono l'oggetto di studio della Ricerca operativa, disciplina nata essenzialmente per esigenze belliche negli anni immediatamente precedenti la seconda guerra mondiale e che ha trovato in seguito il proprio campo di applicazione in settori sempre più vasti e diversi tra loro – e certo non soltanto in quelli bellici!
Il problema che vogliamo analizzare rientra nell'ambito della *programmazione lineare*: si tratta cioè di un problema in cui l'obiettivo è quello di determinare il minimo (*minimizzare*) o di trovare il massimo (*massimizzare*) di una funzione lineare in più variabili (detta *funzione obiettivo*), che a sua volta deve rispettare alcune condizioni (dette *vincoli*), esprimibili mediante equazioni e disequazioni lineari nelle variabili in considerazione.
In altre parole, ottimizzare la funzione obiettivo $f(x_1,\dots,x_n)$ di n variabili equivale a determinare n numeri reali $x_1^*,\dots,x_n^*$ che soddisfano il sistema dei vincoli e rendono minimo o massimo il valore della funzione obiettivo.

Torniamo al problema di partenza, quello della dieta.
Ci proponiamo di scegliere una dieta che soddisfi i requisiti minimi nutrizionali richiesti per rimanere in forma e, perché no, visti i tempi di crisi, spendere il meno possibile.
Cominciamo considerando solo quattro alimenti: pasta, pane, carne e insalata e supponiamo di conoscere il costo medio unitario (in € per 100 grammi), le quantità di fibra, proteine e grassi in essi contenuti, la quantità massima e minima da assumere giornalmente sia di tali nutrienti che degli alimenti in questione (entrambe espresse in grammi).
La tabella seguente riassume tutti i nostri dati:

Alimenti	Fibra	Proteine	Lipidi	Carboidrati	Costo
Pasta	1.5	4.7	0.5	30.3	0.13
Pane	3.1	9	1.9	57.6	0.27
Carne	0	21.5	3.4	0	1.30
Insalata	1.5	1.1	0.1	2.2	0.16
Min nutrienti	10	50	10	50	
Max nutrienti	30	120	80	350	

	Pasta	Pane	Carne	Insalata
Min	60	50	80	100
Max	150	180	350	300

UN PROBLEMA DI PROGRAMMAZIONE LINEARE

Supponiamo di voler determinare due quantità x_1 e x_2 in modo da:

massimizzare $2x_2 - x_1$

con i seguenti vincoli: $4x_1 + 8x_2 \geq -16$

$x_1 + x_2 \leq 8$

$x_1 - x_2 \leq 8$

$x_1 \geq -2$

In questo problema, in cui sono presenti soltanto due variabili, la soluzione può essere ottenuta con un semplice metodo grafico. Prima di tutto si tracciano su un diagramma cartesiano le rette associate a ciascun vincolo, cioè le quattro rette di equazioni:

$4x_1 + 8x_2 = -16$,

$x_1 + x_2 = 8$,

$x_1 - x_2 = 8$,

$x_1 = -2$.

In generale, una disuguaglianza in *n* variabili divide lo spazio *n*-dimensionale in due semispazi: in uno dei due la disuguaglianza è soddisfatta, nell'altro no. La regione ottenuta come intersezione di tutti i semispazi "buoni" sarà l'insieme di tutte le possibili soluzioni, detto *insieme delle soluzioni ammissibili* o *insieme ammissibile*, che a seconda dei casi può essere limitato, illimitato o addirittura vuoto.

La Figura 1 mostra la regione ammissibile del problema considerato, che in questo caso è un poligono. Ciascuna coppia di coordinate di punti interni a esso è detta *soluzione ammissibile*, mentre le coppie di valori che corrispondono ai vertici sono dette *soluzioni ammissibili di base*.

Il teorema fondamentale della programmazione lineare afferma che, sotto opportune ipotesi, una soluzione ottima, cioè un punto che ha per coordinate la particolare coppia di numeri reali che rende massima (o minima) la funzione obiettivo, se esiste, si trova in corrispondenza di uno dei vertici del poligono e non al suo interno. Questo risultato può essere esteso al caso di *n* variabili per problemi di programmazione lineare in forma standard.

In altre parole, una volta individuata la regione ammissibile, il valore massimo (risp. minimo) si ottiene selezionando nel fascio di rette parallele quelle che hanno intersezione non vuota con la regione ammissibile e, successivamente, scegliendo tra queste la retta corrispondente al valore più grande (risp. più piccolo) della funzione obiettivo. I punti di intersezione saranno le soluzioni ottime del problema di programmazione lineare in questione.

Nel caso della Figura 1, le curve di livello della funzione obiettivo sono rette perpendicolari al vettore di coordinate (–1, 2). Il punto ottimo sarà quindi il punto A=(–2, 8), corrispondente al valore ottimo 18.

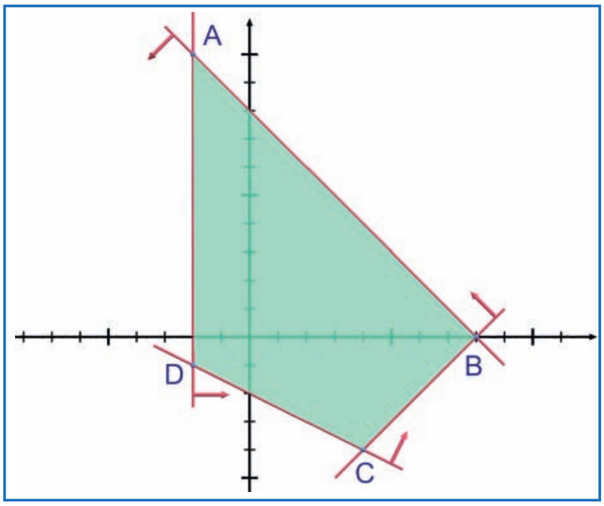

Figura 1. L'insieme delle soluzioni ammissibili.

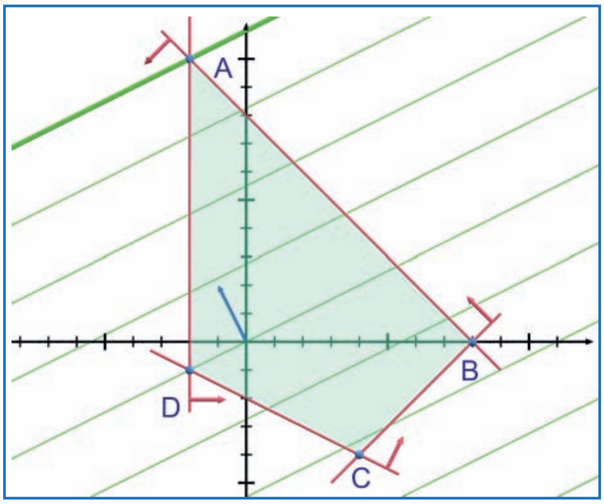

Figura 2. Le curve di livello e la soluzione ottima nell'esempio di Figura 1.

Se indicassimo con x_1 la quantità di pasta, x_2 la quantità di pane, x_3 la quantità di carne, e x_4 la quantità di insalata, considerando che tali quantità non possono assumere valori negativi, otterremmo il seguente problema di programmazione lineare in quattro variabili (e quindi in 4 dimensioni):

minimizzare $0{,}13x_1 + 0{,}27x_2 + 1{,}30x_3 + 0{,}16x_4$

con i seguenti vincoli $10 \leq 1{,}5x_1 + 3{,}1x_2 + 1{,}5x_4 \leq 30$

$50 \leq 4{,}7x_1 + 9x_2 + 21{,}5x_3 + 1{,}1x_4 \leq 120$

$10 \leq 0{,}5x_1 + 1{,}9x_2 + 3{,}4x_3 + 0{,}1\, x_4 \leq 80$

$50 \leq 30{,}3x_1 + 57{,}6x_2 + 2{,}2x_4 \leq 350$

$0{,}6 \leq x_1 \leq 1{,}5$

$0{,}5 \leq x_2 \leq 1{,}8$

$0{,}8 \leq x_3 \leq 3{,}5$

$1 \leq x_4 \leq 3$.

In questo esempio la funzione obiettivo che vogliamo minimizzare rappresenta il costo totale della nostra dieta, calcolato moltiplicando i costi unitari dei vari cibi per la quantità (incognita) di cibo assunto. Inoltre abbiamo 4 vincoli, ciascuno dei quali è relativo a un certo valore nutrizionale. Ad esempio, il primo vincolo impone che la quantità di fibre totali assunte mangiando pasta, pane, carne e insalata sia compresa tra i 10 e i 30 grammi, mentre nei successivi altri vincoli si limitano le quantità di proteine, lipidi e carboidrati.

È possibile dimostrare che il valore ottimo della funzione obiettivo è 3.0718 e $x_1 = 1.5$, $x_2 = 1.8$, $x_3 = 1.67$, $x_4 = 1.44$; dunque la nostra dieta include 150 grammi di pasta, 180 grammi di pane, 167 di carne e 144 di insalata spendendo solo circa 3.07 euro .

In realtà un buon nutrizionista non baserà una dieta equilibrata soltanto su quattro alimenti, bensì su una vasta gamma di essi. In altre parole, se indicassimo con x_1, x_2, ..., x_{n-1}, x_n le quantità di alimenti da acquistare, e quindi da inserire nella nostra dieta, avremmo in generale:

- n alimenti,
- m sostanze nutrienti,
- a_{ij} quantità di sostanza nutriente i presente nell'alimento j,
- c_1, c_2, ..., c_n costi unitari degli alimenti,
- b_1, b_2, ..., b_n quantità giornaliere richieste per ogni sostanza.

Da un punto di vista matematico, tutti i dati a nostra disposizione si possono descrivere tramite il seguente modello, il cui scopo è ancora quello di minimizzare il costo totale della nostra dieta:

$$\begin{array}{ll} \textit{minimizzare} & c_1x_1 + c_2x_2 + \ldots\ldots + c_nx_n \\ \textit{con i seguenti vincoli} & a_{11}x_1 + a_{12}x_2 + \ldots\ldots a_{1n}x_n \geq b_1 \\ & a_{21}x_1 + a_{22}x_2 + \ldots\ldots a_{2n}x_n \geq b_2 \\ & \ldots\ldots \\ & a_{n1}x_1 + a_{n2}x_2 + \ldots\ldots a_{nn}x_n \geq b_n \\ & d_i \leq x_i \leq d'_i,\ i = 1, \ldots\ldots, n, \end{array}$$

dove d_i e d'_i rappresentano la quantità minima e massima del cibo i-simo.

Naturalmente, al problema possono essere aggiunti altri tipi di vincoli: per esempio, vincolo sulle quantità massime giornaliere di sostanze nutrienti o sulla quantità massima di alimenti da assumere, vincoli di *budget*, di preferenza (ma in quest'ultimo caso bisognerebbe considerare, oltre all'obiettivo di minimizzare il costo, anche quello di massimizzare le preferenze).

Un problema di ottimizzazione in n dimensioni è piuttosto complesso. La regione delle soluzioni ammissibili non è più un poligono come nel box di p. 9, bensì un *poliedro* o un *politopo*, l'analogo in dimensione superiore a 2 di un poligono. In tal caso, calcolare il valore della funzione di costo in ogni vertice sarebbe eccessivamente lungo e complicato, visto il numero di vertici possibili. Infatti per un problema di programmazione in forma standard, se fissiamo m vincoli e n variabili, con $n \geq m$, il numero di vertici cresce molto rapidamente (per la precisione, in modo esponenziale) al crescere del numero delle variabili e dei vincoli considerati. Ad esempio, supponendo di avere 6 variabili e 5 vincoli, si può avere un poliedro con un massimo di 6 vertici; se invece il numero di variabili sale a 30, e il numero di vincoli a 20, il numero di vertici può arrivare a oltre 30 milioni!! Per questo il matematico George Dantzig, considerato il padre della programmazione lineare, ideò un algoritmo, detto *metodo del simplesso*, estremamente efficiente nella

				MIN MAX	2000 2250	0 300	0 65	0 2400	0 300
MAX					Calorie	Colesterolo	Grasso	Sodio	Carboidrati
0	0	Broccoli congelati	€ 0,16	250 grammi	73,8	0	0,8	68,2	13,6
0	10	Carote, crude	€ 0,07	1/2 porzione a pezzi	23.7	0	0,1	19,2	5,6
0	10	Sedano, crudo	€ 0,04	1 gambo	6,4	0	0,1	34,8	1,5
0	0	Mais congelato	€ 0,18	1/2 porzione	72,2	0	0,6	2,5	17,1
0	10	Lattuga. iceberg, cruda	€ 0,02	1 porzione	2,6	0	0	1,8	0,4
0	0	Peperone, crudo	€ 0,53	1 peperone	20	0	0,1	1,5	4,8
0	0	Patate, al forno	€ 0,06	1/2 porzione	171,5	0	0,2	15,2	39,9
0	0	Tofu (formaggio di soia)	€ 0,31	1/4 di forma	88,2	0	5,5	8,1	2,2
0	10	Pollo arrostito	€ 0,84	1 porzione media di pollo	277,4	129,9	10,8	125,6	0
0	0	Spaghetti con sugo	€ 0,78	1 1/2 porzione	358,2	0	12,3	1237,1	58,3
0	10	Pomodoro, rosso, maturo, crudo	€ 0,27	1 pomodoro, 2 grandi 3 piccoli	25,8	0	0,4	11,1	5,7
0	10	Mela, cruda con buccia	€ 0,24	1 quantità	81,4	0	0,5	0	21
0	0	Banana	€ 0,15	1 quantità	104,9	0	0,5	1,1	26,7
0	10	Uva	€ 0,32	10 chicchi	15,1	0	0,1	0,5	4,1
0	0	Kiwi crudo fresco	€ 0,49	1 quantità	46,4	0	0,3	3,8	11,3
0	0	Arance	€ 0,15	1 quantità	61,6	0	0,2	0	15,4
0	0	Pane di grano	€ 0,16	24 grammi	78	0	0,5	151,4	15,1
0	0	Pane con cereali	€ 0,05	60 grammi	65	0	1	134,5	12,4

Figura 3. Dati del Dipartimento dell'Agricoltura degli Stati Uniti relativi a cibi e nutrienti disponibili su www.ars.usda.gov

Broccoli congelati		0
Carote, crude	x	0,22281
Sedano, crudo		0
Mais congelato		0
Lattuga, Iceberg, cruda	x	0
Pepperone, crudo		0
Patate, al forno	x	3,132286
Tofu (formaggio di soia)		0
Pollo arrostito	x	0,18315
Spaghetti con sugo	x	0
Pomodoro, rosso, maturo, (	X	0
Mela, cruda con buccia	X	0
Banana	x	0
Uva	x	0
Kiwi crudo fresco	x	0
Arance	x	0
Pane di grano	x	0
Pane con cereali	x	10
Pane bianco	x	1,472897
Farinata d'avena cotta		0
Torta di mela	X	0
Biscotti al cioccolato	X	3,365873
Burro	X	2,532418

Formaggio duro	X	2,728064
Latte intero con 3,3% di grassi		0
Latte parzialmente scremato 2		0
Latte scremato		0
Uovo affogato		0
Uova strapazzate		0
Tacchino	X	0
Wurstel		0
Prosciutto cotto a fette	X	0
Salsiccia affumicata		0
Granola		0
Cheerios		0
Cornflakes		0
Raisin Brn, Kellg'S		0
Rice Krispies		0
Special K		0
Farinata d'avena		0
Malt-O-Meal,Choc		0
Pizza con pepperoni	x	0
Taco (cibo messicano)		0
Hamburger condito		0
Hotdog, liscio		0
Couscous		0

costo		€ 1,86
Riso	X	0
Maccheroni	X	0
Burro di arachidi		0
Maiale		0
Sardine sotto olio		0
Tonno al naturale	X	0
Popcorn		0
Patatine fritte	X	0
Ciambellina salata (tarallini)	X	0
Tortilla		0
Minestra di pollo		0
Pasta con panna e prosciutt	X	0
Zuppa vegetale	X	0
Molluschi		0
Zuppa di pomodoro		0
Molluschi conditi		0
Zuppa di pesce		0
Zuppa di legumi		0

Figura 4. Soluzione di dieta a costo minimo (Network Enabled Optimization System)

pratica. Tale algoritmo permette di ottenere, in un numero finito di passi, una soluzione ottima o di riconoscere che il problema non ha soluzione. Partendo da una soluzione ammissibile di base (corrispondente a un vertice del poliedro) nella regione ammissibile, il metodo procede scegliendo, tra i soli vertici adiacenti a quello scelto, il valore migliore della funzione obiettivo. Dal nuovo vertice trovato si continua poi in modo analogo, visitando solo i vertici adiacenti, finché non si è raggiunta la soluzione ottima.

Sebbene il metodo del simplesso per problemi di programmazione lineare di piccole dimensioni possa essere svolto a mano, quando le dimensioni del problema aumentano diventa sempre più laborioso e complicato. Tuttavia oggi esistono *software* estremamente efficienti, in grado di risolvere in un tempo limitato problemi reali con centinaia di migliaia di variabili e milioni di vincoli.
Anche Excel implementa al suo interno il metodo del simplesso, tramite la componente "Risolutore". Nelle tabelle in Figura 3 e Figura 4 possiamo vedere come tale risolutore risolva il problema relativo ai dati, su cibi e nutrienti, forniti dall'United States Departments of Agricolture (www.ars.usda.gov): più precisamente, nella tabella in Figura 3, vediamo una parte del *database* degli alimenti con tutte le relative quantità di sostanze nutrienti, e nella tabella di figura 4 è possibile vedere la soluzione ottima proposta prendendo in considerazione gli alimenti indicati con la "x" fra tutti quelli disponibili. All'indirizzo web (http://www.neos-guide.org/NEOS/index.php/Diet_Problem_Demo) si può anche utilizzare una versione *online* di questo stesso modello. Nel modello implementato con il Solver di Excel è possibile persino selezionare gli alimenti che più gradiamo ed eliminare quelli che invece proprio non ci vanno giù. Ad analizzare rapidamente le possibili soluzioni e fornire le quantità ottimali e il costo complessivo ci penserà il Risolutore: a noi rimarrà solo da fare la spesa, e concederci persino qualche biscotto al cioccolato!!!

Se questo breve articolo è riuscito nell'intento di suscitare un po' di interesse e curiosità, gli autori, vi invitano ad ampliare e approfondire le idee e i concetti in esso accennati.
Per quanto riguarda la parte teorica potremo indicarvi *Ricerca operativa* di Hillier Frederick S., (ed. McGraw-Hill). Se siete interessati a materiale didattico o divulgativo vi consigliamo:
- http://www.learnaboutor.co.uk/: un sito strettamente collegato alla società inglese di Ricerca operativa per la formazione degli insegnanti e l'allenamento degli studenti alla Ricerca operativa, suddiviso per fasce d'età a partire dagli 11 anni (11-13, 14-16, oltre 16 e studenti universitari) e con contenuti diversi.
- http://www.mindsetproject.org/: una pagina web nata da un progetto americano con l'obiettivo di introdurre la Ricerca operativa in tutte le scuole superiori del North Carolina e del Michigan, che prevede l'insegnamento della matematica utilizzando la scienza delle decisioni e gli strumenti dell'ingegneria. Si focalizza l'attenzione sulla mentalità (*mindset*) della matematica nelle scuole superiori "trasformata" grazie a un diverso approccio dato dalla Ricerca operativa.

Sonia De Cosmis
Laureata magistrale in Matematica e Applicazioni presso l'Università degli studi di Camerino, è attualmente all'ultimo anno di corso di dottorato di ricerca in "*Information Sciences and Complex System*" sempre presso l'Università di Camerino; in particolare si occupa di Robust Support Vector Machines per problemi di classificazione e regressione.
sonia.decosmis@unicam.it

Renato De Leone
Professore ordinario di Ricerca operativa presso l'Università degli Studi di Camerino, è stato Presidente dell'AIRO, Associazione Italiana di Ricerca operativa, dal 2007 al 2011. I suoi interessi vanno dall'ottimizzazione non lineare alla programmazione lineare a larga scala, dall'ottimizzazione nei sistemi di trasporto a modelli e algoritmi per problemi di classificazione.
renato.deleone@unicam.it