

# Ottimizzazione Stocastica

Introduzione ai modelli di programmazione matematica e  
programmazione lineare

Luigi De Giovanni

Dipartimento di Matematica, Università di Padova

# Problema di ottimizzazione

Determinare la migliore configurazione di sistemi complessi sotto condizioni di utilizzo di risorse scarse e/o requisiti di performance.

Esempi:

- Mix ottimo di produzione
- Livelli di sicurezza ottimali delle scorte
- Pianificazione di produzione, schedulazione di processi
- Dimensionamento di servizi (e.g. code e serventi) con vincoli QoS
- Determinazione di percorsi ottimali
- Organizzazione dei flussi di dati in una rete di telecom.
- Pianificazione e gestione operativa di reti di trasporto
- Configurazione di reti di distribuzione elettrica
- *etc. etc. etc.*

# Modelli di programmazione matematica

**Descrivono le caratteristiche della soluzione ottima di un problema di ottimizzazione attraverso relazioni matematiche.**

Elementi di un modello di programmazione matematica:

- **Insiemi:** raggruppano gli elementi del sistema;
- **Parametri:** dati del problema, rappresentano delle quantità “note” (in qualche misura) che dipendono dai diversi elementi del sistema;
- **Variabili decisionali o di controllo:** grandezze da determinare (assimilabili a delle incognite) sulle quali possiamo agire per determinare diverse soluzioni alternative del problema;
- **Vincoli:** relazioni matematiche che descrivono le condizioni di ammissibilità delle soluzioni. Discriminano combinazioni di valori delle variabili decisionali che rappresentano soluzioni accettabili;
- **Funzione obiettivo:** quantità da massimizzare o minimizzare, espressa come funzione delle variabili decisionali.

## Un semplice esempio: problema del coltivatore

Un coltivatore ha a disposizione 11 ettari di terreno da coltivare a lattuga o a patate. Le risorse a sua disposizione, oltre al terreno, sono: 70 kg di semi di lattuga, 18 t di tuberi, 145 t di concime. Supponendo che il mercato sia in grado di assorbire tutta la produzione e che i prezzi siano stabili, la resa stimata per la coltivazione di lattuga è di 3000 €/ettaro e quella delle patate è di 5000 €/ettaro. L'assorbimento delle risorse per ogni tipo di coltivazione è di 7 kg di semi e 10 t di concime per ettaro di lattuga, e 3 t di tuberi e 20 t di concime per le patate. Stabilire quanto terreno destinare a lattuga e quanto a patate in modo da massimizzare la resa economica e sfruttando al meglio le risorse disponibili.

# Un semplice esempio: modello matematico

- **Variabili decisionali:** (cosa bisogna decidere?)

$x_L$ : quantità di ettari da coltivare a lattuga

$x_P$ : quantità di ettari da coltivare a patate

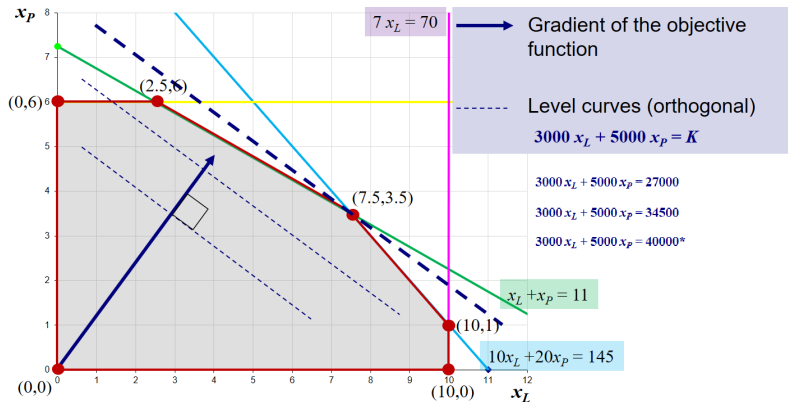
- **Funzione obiettivo:**

$$\max \quad 3000x_L + 5000x_P$$

- **Vincoli:**

$$\begin{array}{rclcl} x_L + x_P & \leq & 11 & \text{(ettari disponibili)} \\ 7x_L & \leq & 70 & \text{(semi disponibili)} \\ & 3x_P & \leq & 18 \text{ (tuberi disponibili)} \\ 10x_L + 20x_P & \leq & 145 & \text{(concime disponibile)} \\ x_L, x_P & \geq & 0 & \text{(dominio)} \end{array}$$

# Un semplice esempio: soluzione



Soluzione semplificata grazie a **relazioni lineari** e variabili continue

Problema più complesso con variabili intere e/o relazioni non lineari: metodi appositi, sia esatti sia euristici, per Programmazione Lineare Mista Intera e Programmazione Non Lineare (e.g., Branch-and-bound, gradient-based, ricerca locale etc.)

# Modelli di programmazione matematica

$$\begin{array}{lll} \min(\max) & f(x) & \\ \text{s.t.} & g_i(x) = b_i & (i = 1 \dots k) \\ & g_i(x) \leq b_i & (i = k + 1 \dots k') \\ & g_i(x) \geq b_i & (i = k' + 1 \dots m) \\ & x_i \in \mathbb{R} \text{ (o } \mathbb{Z} \text{ o } \{0,1\}) & (i = 1 \dots n) \end{array}$$

- $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$  è un vettore (colonna) di  $n$  variabili **reali o intere o binarie**
- $f$  e  $g_i$  sono funzioni  $\mathbb{R}^n \rightarrow \mathbb{R}$
- $b_i \in \mathbb{R}$

# Uso dei modelli matematici

- Il modello matematico ha **valenza descrittiva**: dichiara le caratteristiche di una soluzione ottima del problema. Una soluzione ottima è una combinazione di valori che, assegnati alle variabili decisionali, permettono di rispettare tutti i vincoli (compresi quelli di dominio) e di ottimizzare il valore della funzione obiettivo.
- Il modello matematico ha **valenza operativa**: esistono **algoritmi “standard”** in grado di risolvere il modello, ossia trovare una combinazione di valori ammissibili per le variabili che ottimizza il valore della funzione obiettivo.
- Gli algoritmi **più efficienti considerano  $f$  e  $g_i$  lineari** (compromesso, se possibile, tra potenza espressiva e velocità di soluzione).
- Gli algoritmi di soluzione sono disponibili in **software di ottimizzazione** (Mathematical Programming Solver) in grado di accettare la descrizione del modello e calcolare una soluzione ottima.
- Esempi: linguaggio AMPL, solver Cplex, *Risolutore* di Excel ... e **molti altri**



# Modelli di Programmazione Lineare (PL)

$f$  e  $g_i$  sono funzioni **lineari** di  $x$

$$\begin{array}{ll}\min(\max) & c_1x_1 + c_2x_2 + \dots + c_nx_n \\ \text{s.t.} & a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i \quad (i = 1 \dots k) \\ & a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n \leq b_i \quad (i = k + 1 \dots k') \\ & a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n \geq b_i \quad (i = k' + 1 \dots m) \\ & x_i \in \mathbb{R} \text{ (o } \mathbb{Z} \text{ o } \{0, 1\}) \quad (i = 1 \dots n)\end{array}$$

Classificazione dei modelli di programmazione lineare:

- **modelli di Programmazione Lineare (in senso stretto) PL**: tutte le variabili devono assumere valori reali;
- **modelli di Programmazione Lineare Intera (PLI)** tutte le variabili devono assumere valori interi;
- **modelli di Programmazione Lineare Intera Mista (PLIM)** alcune variabili devono assumere valori reali e altre valori interi.

## “Difficoltà” dei modelli: pillole

La possibilità di risolvere un modello di programmazione matematica in modo più o meno efficiente (velocità di soluzione) dipende dalle caratteristiche del modello stesso

- Modelli di Programmazione Lineare (**PL**)
  - vincoli e funzione obiettivo lineari, variabili continue
  - sono noti algoritmi “efficienti” (di complessità computazionale “polinomiale”)
- Modelli di Programmazione Lineare Intera (**PLI**) e Modelli di Programmazione Lineare Intera Mista (**PLIM**)
  - vincoli e funzione obiettivo lineari, tutte (PLI) o alcune (PLIM) variabili sono intere (le altre continue)
  - in generale, NON sono noti algoritmi “efficienti”
- Modelli di Programmazione Non Lineare (**PNL**)
  - (alcuni) vincoli e/o la funzione obiettivo non sono lineari, variabili continue
  - in generale, NON sono noti algoritmi “efficienti”
- Modelli di Programmazione Non Lineare Intera (**PNLI**) e Modelli di Programmazione NON Lineare Intera Mista (**PNLIM**)
  - (alcuni) vincoli e/o funzione obiettivo non sono lineari, tutte (PNLI) o alcune (PNLIM) variabili sono intere (le altre continue)
  - in generale, NON sono noti algoritmi “efficienti”

In generale, in prima approssimazione, il livello di difficoltà aumenta con l'aumentare delle componenti intere e non lineari.

# Modello PL di un problemi di produzione

## Example

A perfume firm produces two new items by mixing three essences: rose, lily and violet. For each decaliter of perfume *one*, it is necessary to use 1.5 liters of rose, 1 liter of lily and 0.3 liters of violet. For each decaliter of perfume *two*, it is necessary to use 1 liter of rose, 1 liter of lily and 0.5 liters of violet. 27, 21 and 9 liters of rose, lily and violet (respectively) are available in stock. The company makes a profit of 130 euros for each decaliter of perfume *one* sold, and a profit of 100 euros for each decaliter of perfume *two* sold. The problem is to determine the optimal amount of the two perfumes that should be produced.

max	$130 x_{one}$	+	$100 x_{two}$		objective function
s.t.	$1.5 x_{one}$	+	$x_{two}$	$\leq 27$	availability of rose
	$x_{one}$	+	$x_{two}$	$\leq 21$	availability of lily
	$0.3 x_{one}$	+	$0.5 x_{two}$	$\leq 9$	availability of violet
	$x_{one}$	,	$x_{two}$	$\geq 0$	domains of the variables

# Schema generale: mix ottimo di produzione

- **set**  $I$ : resource set       $I = \{rose, lily, violet\}$
- **set**  $J$ : product set       $J = \{one, two\}$
- **parameters**  $D_i$ : availability of resource  $i \in I$       e.g.  $D_{rose} = 27$
- **parameters**  $P_j$ : unit profit for product  $j \in J$       e.g.  $P_{one} = 130$
- **parameters**  $Q_{ij}$ : amount of resource  $i \in I$  required for each unit of product  $j \in J$       e.g.  $Q_{rose\ one} = 1.5, Q_{lily\ two} = 1$
- **variables**  $x_j$ : amount of product  $j \in J$        $x_{one}, x_{two}$

$$\begin{aligned} \max \quad & \sum_{j \in J} P_j x_j \\ \text{s.t.} \quad & \sum_{j \in J} Q_{ij} x_j \leq D_i \quad \forall i \in I \\ & x_j \in \mathbb{R}_+ \quad [ \mathbb{Z}_+ \quad | \quad \{0, 1\} ] \quad \forall j \in J \end{aligned}$$

# Il problema della dieta

## Example

We need to prepare a diet that supplies at least 20 mg of proteins. 30 mg of iron and 10 mg of calcium. We have the opportunity of buying vegetables (containing 5 mg/kg of proteins, 6 mg/Kg of iron e 5 mg/Kg of calcium, cost 4 E/Kg), meat (15 mg/kg of proteins, 10 mg/Kg of iron e 3 mg/Kg of calcium, cost 10 E/Kg) and fruits (4 mg/kg of proteins, 5 mg/Kg of iron e 12 mg/Kg of calcium, cost 7 E/Kg). We want to determine the minimum cost diet.

$$\begin{array}{llllll} \min & 4x_V & + & 10x_M & + & 7x_F & \text{cost} \\ \text{s.t.} & 5x_V & + & 15x_M & + & 4x_F & \geq 20 & \text{proteins} \\ & 6x_V & + & 10x_M & + & 5x_F & \geq 30 & \text{iron} \\ & 5x_V & + & 3x_M & + & 12x_F & \geq 10 & \text{calcium} \\ & x_V & , & x_M & , & x_F & \geq 0 & \text{domains of the variables} \end{array}$$

# Schema generale: copertura di costo minimo

- **set**  $I$ : available resources  $I = \{V, M, F\}$
- **set**  $J$ : request set  $J = \{\textit{proteins}, \textit{iron}, \textit{calcium}\}$
- **parameters**  $C_i$ : unit cost of resource  $i \in I$
- **parameters**  $R_j$ : requested amount of  $j \in J$
- **parameters**  $A_{ij}$ : amount of request  $j \in J$  satisfied by one unit of resource  $i \in I$
- **variables**  $x_i$ : amount of resource  $i \in I$

$$\begin{aligned} \min \quad & \sum_{i \in I} C_i x_i \\ \text{s.t.} \quad & \sum_{i \in I} A_{ij} x_i \geq R_j \quad \forall j \in J \\ & x_i \in \mathbb{R}_+ [\mathbb{Z}_+ \mid \{0, 1\}] \quad \forall i \in I \end{aligned}$$

# Un problema di trasporti

## Example

A company produces refrigerators in three different factories (A, B and C) and need to move them to four stores (1, 2, 3, 4). The production of factories A, B and C is 50, 70 and 20 units, respectively. Stores 1, 2, 3 and 4 require 10, 60, 30 e 40 units, respectively. The costs in Euros to move one refrigerator from a factory to stores 1, 2, 3 and 4 are the following:

from A: 6, 8, 3, 4

from B: 2, 3, 1, 3

from C: 2, 4, 6, 5

The company asks us to formulate a minimum cost transportation plan.

# Scema generale: trasporto

- **set**  $I$ : origins      **factories**  $I = \{A, B, C\}$
- **set**  $J$ : destinations      **stores**  $J = \{1, 2, 3, 4\}$
- **parameters**  $O_i$ : capacity of origin  $i \in I$       **factory production**
- **parameters**  $D_j$ : request of destination  $j \in J$       **store request**
- **parameters**  $C_{ij}$ : unit transp. cost from origin  $i \in I$  to destination  $j \in J$
- **variables**  $x_{ij}$ : amount to be transported from  $i \in I$  to  $j \in J$

$$\min \sum_{i \in I} \sum_{j \in J} C_{ij} x_{ij}$$

s.t.

$$\sum_{i \in I} x_{ij} \geq D_j \quad \forall j \in J$$

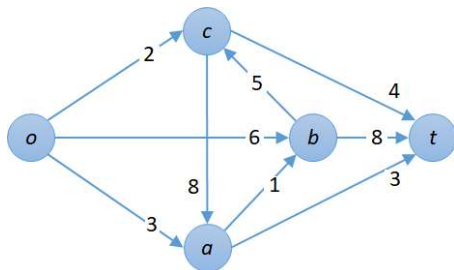
$$\sum_{j \in J} x_{ij} \leq O_i \quad \forall i \in I$$

$$x_{ij} \in \mathbb{R}_+ [\mathbb{Z}_+ \mid \{0, 1\}] \quad \forall i \in I, j \in J$$



# Un problema di cammino minimo

Determinare il cammino minimo da O a T nel seguente grafo.



- cammino: selezione opportuna di archi
- variabile:  $x_{ij} \in \{0, 1\}$   
vale 1 se l'arco è selezionato, 0 altrimenti

$$\begin{aligned} \min \quad & 3x_{oa} + 6x_{ob} + 2x_{oc} + 1x_{ab} + 3x_{at} + 5x_{bc} + 8x_{bt} + 8x_{ca} + 4x_{ct} \\ \text{s.t.} \quad & x_{oa} + x_{ob} + x_{oc} = 1 \\ & x_{at} + x_{bt} + x_{ct} = 1 \\ & x_{oc} + x_{bc} - (x_{ca} + x_{ct}) = 0 \\ & x_{ob} + x_{ab} - (x_{bc} + x_{bt}) = 0 \\ & x_{oa} + x_{ca} - (x_{ab} + x_{at}) = 0 \\ & x_{ij} \in \{0, 1\} \end{aligned}$$



## Problema del cammino minimo: generalizzazione

Dato un grafo  $G = (N, A)$  con un costo  $c_{ij}$  associato a ogni arco  $(i, j) \in A$ , determinare un cammino di costo minimo da un nodo sorgente  $s \in N$  a un nodo destinazione  $d \in N$

Variabili decisionali:  $x_{ij} = \begin{cases} 1, & \text{l'arco } (i, j) \text{ è sul cammino minimo;} \\ 0, & \text{altrimenti.} \end{cases}$

$$\min \quad \sum_{(i,j) \in A} c_{ij} x_{ij}$$

s.t.

$$\sum_{(i,v) \in A} x_{iv} - \sum_{(v,j) \in A} x_{vj} = \begin{cases} 0, & v \in N \setminus \{s, d\}. \\ -1, & v = s; \\ +1, & v = d; \end{cases}$$

$$x_{ij} \in \{0, 1\} \equiv \mathbb{R}_+ \quad \forall (i, j) \in A$$

# Programm. matematica e condizioni di incertezza o rischio

## Example

Nel problema del coltivatore, considerare l'incertezza legata alla resa economica (funzione obiettivo - FO), rappresentata attraverso 3 scenari. La resa unitaria di lattuga e patata è: nello scenario 1, 3000 e 5000 €; nello scenario 2, 2000 e 8000 €; nello scenario 3, 4000 e 2000 €.

- a. Scegliere la configurazione che massimizza il minimo della resa tra tutti gli scenari.
- b. Data probabilità a priori per gli scenari 1, 2 e 3 pari risp. a 0.4, 0.4 e 0.2, determinare la configurazione che massimizza la resa media.

Utilizziamo il **modello matematico** per rappresentare implicitamente tutte le **soluzioni alternative** (regione ammissibile del modello), selezionando quella che risponde al criterio in esame.

Nota: l'analisi decisionale presuppone un insieme dato di alternative



## Caso FO.a: max - min

$$\max \min \{3x_L + 5x_P ; 2x_L + 8x_P ; 5x_L + 4x_P\}$$

s.t.

$$\left. \begin{array}{rclcl} x_L + x_P & \leq & 11 & \text{(ettari)} \\ 7x_L & \leq & 70 & \text{(semi)} \\ & 3x_P & \leq & 18 \text{ (tuberi)} \\ 10x_L + 20x_P & \leq & 145 & \text{(concime)} \\ x_L, x_P & \geq & 0 & \text{(dominio)} \end{array} \right\} F \text{ (regione ammissibile)}$$

Attenzione: la funzione obiettivo può essere resa lineare come segue:

$$\begin{array}{ll} \max & y \\ \text{s.t.} & y \leq 3x_L + 5x_P \\ & y \leq 2x_L + 8x_P \\ & y \leq 4x_L + 2x_P \\ & (x_L, x_P) \in F \\ & y \in \mathbb{R} \end{array}$$

## Caso FO.b: max Valore Atteso

$$\max 0.4(3x_L + 5x_P) + 0.4(2x_L + 8x_P) + 0.2(4x_L + 2x_P)$$

s.t.

$$\left. \begin{array}{rclcl} x_L + x_P & \leq & 11 & \text{(ettari)} \\ 7x_L & \leq & 70 & \text{(semi)} \\ & 3x_P & \leq & 18 \text{ (tuberi)} \\ 10x_L + 20x_P & \leq & 145 & \text{(concime)} \\ x_L, x_P & \geq & 0 & \text{(dominio)} \end{array} \right\} F \text{ (regione ammissibile)}$$





# Programm. matematica e condizioni di incertezza o rischio

## Example

Nel problema del coltivatore, considerare l'incertezza legata alle risorse disponibili (regione ammissibile - RA), rappresentata attraverso 3 scenari. La disponibilità di tuberi e di concime è: nello scenario 1, 18 e 145; nello scenario 2, 16 e 165; nello scenario 3, 25 e 135.

- a. Scegliere la configurazione che massimizza la resa e rimane ammissibile in tutti (caso a1) o in almeno 2 (caso a2) scenari
- b. Data probabilità a priori per gli scenari 1, 2 e 3 pari risp. a 0.4, 0.4 e 0.2, determinare la configurazione che massimizza la resa, rimanendo ammissibile con probabilità pari o superiore al 50%.

Utilizziamo il **modello matematico** per rappresentare implicitamente tutte le **soluzioni alternative** che rimangono ammissibili secondo il criterio in esame, selezionando quella migliore.

Nota: l'analisi decisionale assume un insieme dato di alternative e che le alternative date siano ammissibili in ogni scenario (o "recuperabili" con costi di inclusi nella valutazione)



## Caso RA.a1: robustezza assoluta

Le soluzioni devono soddisfare i vincoli di tutti gli scenari.

$$\max 3x_L + 5x_P$$

s.t.

$$\begin{array}{rcllcl} x_L & + & x_P & \leq & 11 & \text{(ettari)} \\ 7x_L & & & \leq & 70 & \text{(semi)} \\ & & 3x_P & \leq & 18 & \text{(scen. 1 - tuberi)*} \\ 10x_L & + & 20x_P & \leq & 145 & \text{(scen. 1 - concime)*} \\ & & 3x_P & \leq & 16 & \text{(scen. 2 - tuberi)} \\ 10x_L & + & 20x_P & \leq & 165 & \text{(scen. 2 - concime)*} \\ & & 3x_P & \leq & 25 & \text{(scen. 3 - tuberi)*} \\ 10x_L & + & 20x_P & \leq & 135 & \text{(scen. 3 - concime)} \\ x_L & , & x_P & \geq & 0 & \text{(dominio)} \end{array}$$

Nota (\*): possiamo eliminare alcuni vincoli ridondanti



## Caso RA.a2: livello di robustezza parametrico

Introduciamo variabili binarie  $z_s$  che indicano se decidiamo di soddisfare i vincoli dello scenario  $s$  (valore 1) o meno (valore 0).

$$\max 3x_L + 5x_P$$

s.t.

$$\begin{array}{llllll} x_L & + & x_P & \leq & 11 & \text{(ettari)} \\ 7x_L & & & \leq & 70 & \text{(semi)} \\ & & 3x_P & \leq & 18 + 1000(1 - z_1) & \text{(scen.1-tuberi)} \\ 10x_L & + & 20x_P & \leq & 145 + 1000(1 - z_1) & \text{(scen.1-concime)} \\ & & 3x_P & \leq & 16 + 1000(1 - z_2) & \text{(scen.2-tuberi)} \\ 10x_L & + & 20x_P & \leq & 165 + 1000(1 - z_2) & \text{(scen.2-concime)} \\ & & 3x_P & \leq & 25 + 1000(1 - z_3) & \text{(scen.3-tuberi)} \\ 10x_L & + & 20x_P & \leq & 135 + 1000(1 - z_3) & \text{(scen.3-concime)} \\ z_1 + z_2 & + & z_3 & \geq & 2 & \text{(almeno 2 scenari)} \\ x_L & , & x_P & \geq & 0 & \text{(dominio)} \\ z_1, z_2 & , & z_3 & \in & \{0, 1\} & \end{array}$$

## Caso RA.b: probabilità di ammissibilità

Sfruttiamo le variabili binarie  $z_s$  e il fatto che, *in questo caso*:

$$\mathbb{P}(\text{sol. ammissibile}) = \sum_s p(s) z_s.$$

$$\max 3x_L + 5x_P$$

s.t.

$$\begin{array}{llllll} x_L & + & x_P & \leq & 11 & \text{(ettari)} \\ 7x_L & & & \leq & 70 & \text{(semi)} \\ & & 3x_P & \leq & 18 + 1000(1 - z_1) & \text{(scen.1-tuberi)} \\ 10x_L & + & 20x_P & \leq & 145 + 1000(1 - z_1) & \text{(scen.1-concime)} \\ & & 3x_P & \leq & 16 + 1000(1 - z_2) & \text{(scen.2-tuberi)} \\ 10x_L & + & 20x_P & \leq & 165 + 1000(1 - z_2) & \text{(scen.2-concime)} \\ & & 3x_P & \leq & 25 + 1000(1 - z_3) & \text{(scen.3-tuberi)} \\ 10x_L & + & 20x_P & \leq & 135 + 1000(1 - z_3) & \text{(scen.3-concime)} \\ 0.4 z_1 + 0.4 z_2 & + & 0.2 z_3 & \geq & 0.5 & \text{(prob. ammissibile)} \\ x_L & , & x_P & \geq & 0 & \text{(dominio)} \\ z_1, z_2 & , & z_3 & \in & \{0, 1\} & \end{array}$$

## In sintesi...

- *Ottimizzare*: costruire la migliore alternativa (non solo confrontare alternative tra loro)
- In condizioni di incertezza o rischio possiamo avere:
  - diverse modalità di **rappresentazione dell'incertezza** (oltre agli scenari fin qui utilizzati come esempio introduttivo)
  - diversi **criteri di robustezza e/o ottimalità** (soluzioni robuste in senso assoluto o relativo [ottimizzazione robusta], soluzioni ottime in condizioni di rischio [ottimizzazione stocastica] **e altri ...**)
    - ▶ come definisco le soluzioni “accettabili” in condizioni di incert / rischio?
    - ▶ come definisco le soluzioni “ottime” in condizioni di incert./rischio?
- La programmazione matematica è uno strumento (non l'unico) che permette di considerare diverse modalità e diversi criteri.

Nota: I modelli presentati sono implementati in AMPL e disponibili nella cartella *modelliampl* (vedi anche la risorsa “Sintassi base di AMPL”). Gli studenti interessanti trovano ulteriori esempi e approfondimenti sulla modellazione in programmazione lineare nel file *modellazionePL.pdf* (in lettura)











