

6.215/6.255J/15.093J/IDS.200J Optimization Methods

Lecture 12: Data Uncertainty in Linear Optimization

October 19, 2021

Today's Lecture

Outline

- dealing with data uncertainty; motivation
- stochastic approach
- robust approach

Linear optimization under data uncertainty

Motivation (example from Ben-Tal 2009)

- Consider PILOT4, a linear optimization problem from the NETLIB benchmark library (<http://www.netlib.org/lp/data/>) with $n = 1000$ variables and $m = 410$ constraints
- One of the constraints (#372) from PILOT4:

$$\begin{aligned} a_{372}^T x = & -15.79081x_{826} - 8.598819x_{827} - 1.88789x_{828} - 1.362417x_{829} - 1.526049x_{830} \\ & - 0.031883x_{849} - 28.725555x_{850} - 10.792065x_{851} - 0.19004x_{852} - 2.757176x_{853} \\ & - 12.290832x_{854} + 717.562256x_{855} - 0.057865x_{856} - 3.785417x_{857} - 78.30661x_{858} \\ & - 122.163055x_{859} - 6.46609x_{860} - 0.48371x_{861} - 0.615264x_{862} - 1.353783x_{863} \\ & - 84.644257x_{864} - 122.459045x_{865} - 43.15593x_{866} - 1.712592x_{870} - 0.401597x_{871} \\ & + 1 \cdot x_{880} - 0.946049x_{898} - 0.946049x_{916} \geq b_{372} = 23.387405 \end{aligned}$$

Linear optimization under data uncertainty

Motivation (example from Ben-Tal 2009)

- Consider PILOT4, a linear optimization problem from the NETLIB benchmark library (<http://www.netlib.org/lp/data/>) with $n = 1000$ variables and $m = 410$ constraints
- One of the constraints (#372) from PILOT4:

$$\begin{aligned} a_{372}^T x = & -15.79081x_{826} - 8.598819x_{827} - 1.88789x_{828} - 1.362417x_{829} - 1.526049x_{830} \\ & - 0.031883x_{849} - 28.725555x_{850} - 10.792065x_{851} - 0.19004x_{852} - 2.757176x_{853} \\ & - 12.290832x_{854} + 717.562256x_{855} - 0.057865x_{856} - 3.785417x_{857} - 78.30661x_{858} \\ & - 122.163055x_{859} - 6.46609x_{860} - 0.48371x_{861} - 0.615264x_{862} - 1.353783x_{863} \\ & - 84.644257x_{864} - 122.459045x_{865} - 43.15593x_{866} - 1.712592x_{870} - 0.401597x_{871} \\ & + 1 \cdot x_{880} - 0.946049x_{898} - 0.946049x_{916} \geq b_{372} = 23.387405 \end{aligned}$$

- Many numbers here such as -8.598819 are potentially inaccurate
- Numbers such as 1 are probably exact

Linear optimization under data uncertainty

Motivation (example from Ben-Tal 2009)

- The related *nonzero* optimal variables reported by the solver CPLEX are:

$$\begin{array}{ll} x_{826}^* = 255.6112787181108, & x_{827}^* = 6240.488912232100, \\ x_{828}^* = 3624.613324098961, & x_{829}^* = 18.20205065283259, \\ x_{849}^* = 174397.0389573037, & x_{870}^* = 14250.00176680900, \\ x_{871}^* = 25910.00731692178, & x_{880}^* = 104958.3199274139. \end{array}$$

- Within machine precision x^* makes the previous constraints active (an equality).
- Suppose the coefficients in \mathbf{a}_{372} are 0.1% accurate approximations of their true values. Will the above solution still be feasible?

Linear optimization under data uncertainty

Motivation (example from Ben-Tal 2009)

- The related *nonzero* optimal variables reported by the solver CPLEX are:

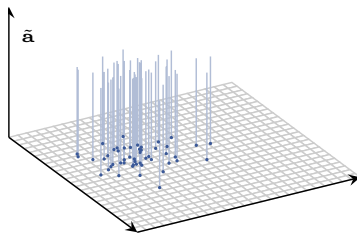
$$\begin{aligned}x_{826}^* &= 255.6112787181108, & x_{827}^* &= 6240.488912232100, \\x_{828}^* &= 3624.613324098961, & x_{829}^* &= 18.20205065283259, \\x_{849}^* &= 174397.0389573037, & x_{870}^* &= 14250.00176680900, \\x_{871}^* &= 25910.00731692178, & x_{880}^* &= 104958.3199274139.\end{aligned}$$

- Within machine precision \mathbf{x}^* makes the previous constraints active (an equality).
- Suppose the coefficients in \mathbf{a}_{372} are 0.1% accurate approximations of their true values. Will the above solution still be feasible?
- Worst-case infeasibility : $\tilde{\mathbf{a}}_{372}^T \mathbf{x}^* - b_{372} = -104.9 \ll 0$
A violation of 448.5% of the right hand side!
- Small perturbations make our “optimal” solution \mathbf{x}^* practically meaningless.

Linear optimization under data uncertainty

Philosophical considerations

- probability distributions are used to model uncertainty but are rarely observed in practice
- data is however observed in practice



- can this provide alternate model of uncertainty with tractable formulations?

Linear optimization under data uncertainty

Describing uncertainty

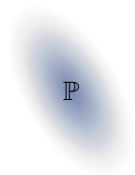
Two fundamentally different approaches to describe an **uncertain quantity** $\tilde{\mathbf{a}}$.

(1) probabilistic description

- a distribution \mathbb{P} describes all possible outcomes

$$\tilde{\mathbf{a}} \sim \mathbb{P}$$

- directly related to *stochastic optimization*

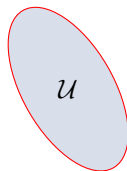


(2) set description

- a set \mathcal{U} describes all possible outcomes

$$\tilde{\mathbf{a}} \in \mathcal{U}$$

- directly related to *robust optimization*



Linear optimization under data uncertainty

A (very short) overview!

- deterministic linear optimization:

$$\min_{\mathbf{x} \in \mathcal{X}} \mathbf{c}^T \mathbf{x} \quad \text{s.t.} \quad \mathbf{A} \mathbf{x} \leq \mathbf{b}$$

- linear optimization under uncertainty $\tilde{\mathbf{z}} \in \mathcal{W}$:

$$\min_{\mathbf{x} \in \mathcal{X}} \mathbf{c}(\tilde{\mathbf{z}})^T \mathbf{x} \quad \text{s.t.} \quad \mathbf{A}(\tilde{\mathbf{z}}) \mathbf{x} \leq \mathbf{b}(\tilde{\mathbf{z}})$$

- without loss of generality, we can always assume that data uncertainty affects only \mathbf{A} and \mathbf{c} , but not the vector \mathbf{b} .
- probability distributions available: stochastic optimization (Prékopa 1995; Birge and Louveaux 1997; etc. ...)
- distributions unavailable:
 - robust optimization (Ben-Tal and Nemirovski 1999; Bertsimas and Sim 2004; Ben-Tal 2009; etc. ...)
 - distributionally robust optimization (Delage and Ye 2010; Wiesemann et al. 2014; etc. ...)

Linear optimization under data uncertainty

The “easy case”, uncertainty in \mathbf{c} only

- linear optimization under uncertainty $\tilde{\mathbf{z}} \in \mathcal{W}$: $\min_{\mathbf{x} \in \mathcal{X}} \mathbf{c}(\tilde{\mathbf{z}})^T \mathbf{x}$ s.t. $\mathbf{Ax} \leq \mathbf{b}$
 - probability distributions available:
replace $\mathbf{c}(\tilde{\mathbf{z}})$ by its expected value, or a linear combination of its expected value and standard deviation, ...
 - distributions unavailable:
 - use lower or upper bounds of the ranges taken by $\mathbf{c}(\tilde{\mathbf{z}})$, ...
- in general, relatively well understood, close to sensitivity analysis, feasibility of a solution is maintained ...

Linear optimization under data uncertainty

Uncertainty in \mathbf{A} only

- linear optimization under uncertainty $\tilde{\mathbf{z}} \in \mathcal{W}$: $\min_{\mathbf{x} \in \mathcal{X}} \mathbf{c}^T \mathbf{x}$ s.t. $\mathbf{A}(\tilde{\mathbf{z}})\mathbf{x} \leq \mathbf{b}$
- probability distributions available:
probabilistic feasibility \Rightarrow chance-constrained optimization

$$\begin{array}{ll} \min & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} & \mathbb{P}(\mathbf{A}(\tilde{\mathbf{z}})\mathbf{x} \leq \mathbf{b}) \geq \delta \\ & \mathbf{x} \in \mathcal{X} \end{array}$$

- comments:
 - need to specify a given \mathbb{P}
 - need to specify a confidence level $\delta \in [0, 1]$
 - in most cases, nonlinear constraints, hard to solve exactly

Linear optimization under data uncertainty

Uncertainty in \mathbf{A} only

- linear optimization under uncertainty $\tilde{\mathbf{z}} \in \mathcal{W}$: $\min_{\mathbf{x} \in \mathcal{X}} \mathbf{c}^T \mathbf{x} \quad \text{s.t.} \quad \mathbf{A}(\tilde{\mathbf{z}}) \mathbf{x} \leq \mathbf{b}$
- probability distributions unavailable:
worst-case feasibility \Rightarrow robust optimization

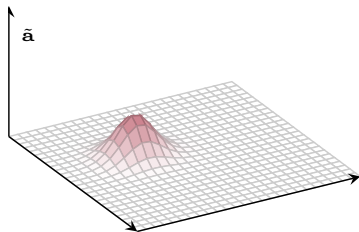
$$\begin{aligned} \min \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}(\mathbf{z}) \mathbf{x} \leq \mathbf{b} \quad \forall \mathbf{z} \in \mathcal{U}(\Gamma) \\ & \mathbf{x} \in \mathcal{X} \end{aligned}$$

- comments:
 - need to specify a family of *uncertainty sets*, $\mathcal{U}(\Gamma)$
 - need to specify a “budget of uncertainty” $\Gamma \in \mathbb{R}_+$, corresponding to the level of uncertainty that must be tolerated.
 - in some important cases, lead to tractable formulations

Linear optimization under data uncertainty

Uncertainty in \mathbf{A} - probabilistic feasibility example

- how feasible is \mathbf{x} in $\tilde{\mathbf{a}}^T \mathbf{x} \leq b$ if the vector $\tilde{\mathbf{a}}$ is uncertain?
- probabilistic feasibility $\Rightarrow \mathbb{P}(\tilde{\mathbf{a}}^T \mathbf{x} \leq b) \geq \delta$ with $\delta \in [0, 1]$.
- assume that $\tilde{\mathbf{a}} \sim \mathbf{N}(\mathbf{a}, \mathbf{\Sigma})$ is a normal random vector.



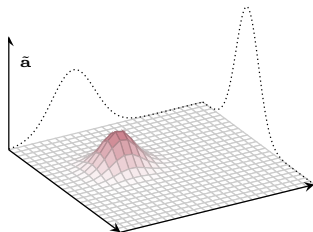
- mean vector: $\mathbf{a} \in \mathbb{R}^n$
- variance-covariance matrix: $\mathbf{\Sigma} \in \mathbb{R}^{n \times n}$.

Linear optimization under data uncertainty

Uncertainty in \mathbf{A} - probabilistic feasibility example

- normal distributions satisfy a projection property:

$$\tilde{\mathbf{a}} \sim N(\mathbf{a}, \Sigma) \implies \tilde{\mathbf{a}}^T \mathbf{x} \sim N(\mathbf{a}^T \mathbf{x}, \mathbf{x}^T \Sigma \mathbf{x})$$



- a specific property of the multivariate normal distribution

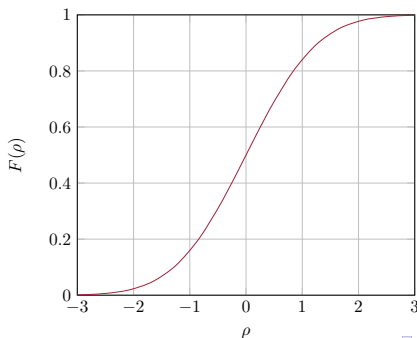
Linear optimization under data uncertainty

Uncertainty in \mathbf{A} - probabilistic feasibility example

- computing the probability for a *single* constraint is then “easy”:

$$\mathbb{P}(\tilde{\mathbf{a}}^T \mathbf{x} \leq b) = \mathbb{P}(\mathbf{N} \leq b) = F\left(\frac{b - \mathbf{a}^T \mathbf{x}}{\sqrt{\mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x}}}\right)$$

where \mathbf{N} is a normal random variable with mean $\mathbf{a}^T \mathbf{x}$ and variance $\mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x}$, and where $F(\cdot)$ is its cumulative distribution function.



Linear optimization under data uncertainty

Uncertainty in \mathbf{A} - probabilistic feasibility example

- stochastic optimization

$$\begin{aligned} \min \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{X} \\ & \mathbb{P} \left(\tilde{\mathbf{a}}^T \mathbf{x} \leq b \right) \geq \delta \end{aligned}$$

- equivalent deterministic formulation

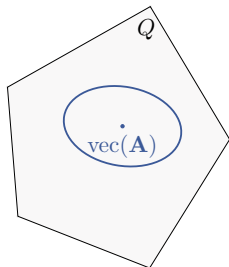
$$\begin{aligned} \min \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{X} \\ & F \left(\frac{b - \mathbf{a}^T \mathbf{x}}{\sqrt{\mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x}}} \right) \geq \delta \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} \min \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{X} \\ & b - \mathbf{a}^T \mathbf{x} \geq F^{-1}(\delta) \cdot \sqrt{\mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x}} \end{aligned}$$

- can be solved using (slower but convex) optimization methods.

Linear optimization under data uncertainty

Uncertainty in \mathbf{A} - probabilistic feasibility example

- multivariate extension would be $\mathbb{P}(\tilde{\mathbf{A}}\mathbf{x} \leq \mathbf{b}) \geq \delta$ with $\text{vec}(\tilde{\mathbf{A}}) \in \mathbb{R}^{m \cdot n}$ a normal random vector.
- verifying feasibility would require us to compute $\mathbb{P}(\text{vec}(\tilde{\mathbf{A}}) \in Q$ where $Q := \{\text{vec}(\mathbf{A}) \in \mathbb{R}^{m \cdot n} : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$)



- integrating in high dimensions is very hard!

Linear optimization under data uncertainty

Uncertainty in \mathbf{A} - probabilistic feasibility - practicality

- even for simple distributions and small problems, calculating the probability of violation of multiple constraints is often difficult.
- simulation is only an approximation and may be very computationally expensive.
- problems get worse as number of variables and constraints grow.
- optimization in these environment is even harder.

Linear optimization under data uncertainty

Uncertainty in \mathbf{A} - worst-case feasibility - robust optimization - a response?

- in practice *robustness* is often as important as *optimality* (sometimes more).
- *motivation 1*: create solutions that are immune to implementation errors and data uncertainty
- *motivation 2*: develop a theory of performance analysis and optimization under uncertainty via optimization that is tractable in high dimensions.
- a remark on *tractability*: we do not mean polynomial solvability; rather the ability to solve problems of the size and in times that are appropriate for the application.

Linear optimization under data uncertainty

Uncertainty in \mathbf{A} - robust optimization

$$\begin{aligned} \min \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \tilde{\mathbf{a}}_1^T \mathbf{x} \leq b_1 \\ & \dots \\ & \tilde{\mathbf{a}}_m^T \mathbf{x} \leq b_m \end{aligned}$$

- uncertain vectors $(\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_m) \in \mathcal{U}$.
- **robust problem formulation**

$$\begin{aligned} \min \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \tilde{\mathbf{a}}_1^T \mathbf{x} \leq b_1 \quad \forall (\tilde{\mathbf{a}}_1, \dots, \tilde{\mathbf{a}}_m) \in \mathcal{U} \\ & \dots \\ & \tilde{\mathbf{a}}_m^T \mathbf{x} \leq b_m \end{aligned}$$

Linear optimization under data uncertainty

Uncertainty in \mathbf{A} - robust optimization - constructing uncertainty sets - illustration

problem:

- variable x_i : complete products $i = [1, 2, 3, 4]$ (e.g., chairs, desks, tables, beds).
- profit: $50x_1 + 40x_2 + 60x_3 + 30x_4$.
- time resource constraint:

$$\tilde{\mathbf{a}}^T \mathbf{x} \leq b \iff 120x_1 + 100x_2 + 180x_3 + 140x_4 \leq 5000$$

robust formulation:

$$\begin{array}{ll} \max & 50x_1 + 40x_2 + 60x_3 + 30x_4 \\ \text{s.t.} & \tilde{\mathbf{a}}^T \mathbf{x} \leq b \end{array}$$

Linear optimization under data uncertainty

Uncertainty in \mathbf{A} - robust optimization - constructing uncertainty sets - illustration

problem:

- variable x_i : complete products $i = [1, 2, 3, 4]$ (e.g., chairs, desks, tables, beds).
- profit: $50x_1 + 40x_2 + 60x_3 + 30x_4$.
- time resource constraint:

$$\tilde{\mathbf{a}}^T \mathbf{x} \leq b \iff 120x_1 + 100x_2 + 180x_3 + 140x_4 \leq 5000$$

the coefficient vector $\tilde{\mathbf{a}} \in \mathcal{U}$ is typically uncertain for new products!

robust formulation:

$$\begin{aligned} \max \quad & 50x_1 + 40x_2 + 60x_3 + 30x_4 \\ \text{s.t.} \quad & \tilde{\mathbf{a}}^T \mathbf{x} \leq b \quad \forall \tilde{\mathbf{a}} \in \mathcal{U} \end{aligned}$$

Linear optimization under data uncertainty

Uncertainty in \mathbf{A} - robust optimization - constructing uncertainty sets - illustration

uncertainty set:

- rule of thumb : we know coefficients up to $\Gamma\%$ inaccuracy
- uncertainty set

$$\mathcal{U} = \left\{ \tilde{\mathbf{a}} \in \mathbb{R}^4 : \begin{array}{l} (1 - \Gamma) \cdot 120 \leq \tilde{a}_1 \leq (1 + \Gamma) \cdot 120, \\ (1 - \Gamma) \cdot 100 \leq \tilde{a}_2 \leq (1 + \Gamma) \cdot 100, \\ (1 - \Gamma) \cdot 180 \leq \tilde{a}_3 \leq (1 + \Gamma) \cdot 180, \\ (1 - \Gamma) \cdot 140 \leq \tilde{a}_4 \leq (1 + \Gamma) \cdot 140. \end{array} \right\}$$

- notice that we insist the solution be feasible for any value of $\tilde{\mathbf{a}}$ in that range.
(in particular, the **worst** values.)
- by varying the value of $\Gamma \geq 0$ we can control level of robustness.

Linear optimization under data uncertainty

Uncertainty in \mathbf{A} - robust optimization - constructing uncertainty sets - illustration

- a fair criticism of the robust approach is that it is “unlikely” that all coefficients take their worst value at once.
- is there a more reasonable uncertainty set?
- suggestions?

Linear optimization under data uncertainty

Uncertainty in **A** - robust optimization - using probability !!

- if coefficients \tilde{a}_i were random variables independent with mean μ and variance σ^2
- central limit theorem: for large n

$$\frac{\tilde{a}_1 + \cdots + \tilde{a}_n - n\mu}{\sigma \cdot \sqrt{n}} \rightarrow N(0, 1).$$

- intuitively, this tells us that sums of **independent** random variables tend to be close to their mean.
- how might we use this to construct an uncertainty set?

$$\mathcal{U} = \left\{ \tilde{\mathbf{a}} : -\Gamma\sqrt{n} \leq \sum_{i=1}^n \frac{\tilde{a}_i - a_i}{\sigma_i} \leq \Gamma\sqrt{n} \right\}$$

Linear optimization under data uncertainty

Uncertainty in \mathbf{A} - robust optimization - typical uncertainty sets

- naive

$$\mathcal{U} = \left\{ \tilde{\mathbf{a}} : \left| \frac{\tilde{a}_i - a_i}{\sigma_i} \right| \leq \Gamma \quad \forall i \in 1, \dots, n \right\}$$

- CLT

$$\mathcal{U} = \left\{ \tilde{\mathbf{a}} : -\Gamma\sqrt{n} \leq \sum_{i=1}^n \frac{\tilde{a}_i - a_i}{\sigma_i} \leq \Gamma\sqrt{n} \right\}$$

- bounded CLT

$$\mathcal{U} = \left\{ \tilde{\mathbf{a}} : \sum_{i=1}^n \left| \frac{\tilde{a}_i - a_i}{\sigma_i} \right| \leq \Gamma\sqrt{n} \right\}$$

- ellipsoidal sets

$$\mathcal{U} = \left\{ \tilde{\mathbf{a}} : \sum_{i=1}^n \frac{(\tilde{a}_i - a_i)^2}{\sigma_i^2} \leq \Gamma\sqrt{n} \right\}$$

- ...

Linear optimization under data uncertainty

Uncertainty in \mathbf{A} - robust optimization - tractability

- consider a special case: robust optimization with row-wise uncertainty

$$\begin{aligned} \min \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \tilde{\mathbf{a}}_i^T \mathbf{x} \leq \mathbf{b}_i \quad \forall \tilde{\mathbf{a}}_i \in U_i, \quad i \in 1, \dots, m. \end{aligned}$$

- primitives: uncertainty sets $U_i, i \in 1, \dots, m$
- problem could be regarded as a linear optimization problem with (possibly infinitely) many constraints \Rightarrow impractical.
- reformulation:

$$\begin{aligned} \min \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \max_{\tilde{\mathbf{a}}_i \in U_i} \{ \tilde{\mathbf{a}}_i^T \mathbf{x} \} \leq \mathbf{b}_i \quad \forall i \in 1, \dots, m. \end{aligned}$$

Linear optimization under data uncertainty

Uncertainty in \mathbf{A} - robust optimization - tractability of the feasibility problem

- suppose that U_i , $i \in 1, \dots, m$ are polyhedral sets.
- given a fixed \mathbf{x} , we can efficiently solve $\forall i \in 1, \dots, m$ the *feasibility problems*:
 - verify whether

$$\max_{\tilde{\mathbf{a}}_i \in U_i} \{\tilde{\mathbf{a}}_i^T \mathbf{x}\} \leq b_i,$$

- or find $\bar{\mathbf{a}}_i \in U_i$ such that

$$\bar{\mathbf{a}}_i^T \mathbf{x} > b_i.$$

- how should we then solve the robust optimization problem ?

Linear optimization under data uncertainty

Uncertainty in \mathbf{A} - robust optimization - robust counterpart

- robust optimization with row-wise uncertainty

$$\begin{aligned} \min \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \max_{\tilde{\mathbf{a}}_i \in U_i} \{\tilde{\mathbf{a}}_i^T \mathbf{x}\} \leq \mathbf{b}_i \quad \forall i \in 1, \dots, m. \end{aligned}$$

- nonempty uncertainty sets $U_i = \{\tilde{\mathbf{a}}_i : \mathbf{D}_i \tilde{\mathbf{a}}_i \leq \mathbf{d}_i\}$ with $\mathbf{D}_i \in \Re^{k_i \times n}$.
- consider the dual of inner maximization problem

$$\begin{aligned} \max \quad & \tilde{\mathbf{a}}_i^T \mathbf{x} & = & \min \quad \mathbf{p}_i^T \mathbf{d}_i \\ \text{s.t.} \quad & \mathbf{D}_i \tilde{\mathbf{a}}_i \leq \mathbf{d}_i, & \text{s.t.} \quad & \mathbf{p}_i \geq 0, \\ & \tilde{\mathbf{a}}_i \text{ free} & & \mathbf{p}_i^T \mathbf{D}_i = \mathbf{x}^T. \end{aligned}$$

- we have strong duality! (note: If primal is unbounded then dual is infeasible and we interpret $+\infty = +\infty$.)

Linear optimization under data uncertainty

Uncertainty in \mathbf{A} - robust optimization - the power of duality !

so robust optimization is a linear optimization itself!

$$\begin{aligned} \min \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{p}_i^T \mathbf{d}_i \leq \mathbf{b}_i \quad \forall i \in 1, \dots, m \\ & \mathbf{p}_i^T \mathbf{D}_i = \mathbf{x}^T \quad \forall i \in 1, \dots, m \\ & \mathbf{p}_i \geq 0 \quad \forall i \in 1, \dots, m \end{aligned}$$

- nominal problem : n variables with m constraints.
- robust problem : $n + \sum_{i=1}^m k_i$ variables with $m + n \cdot m + \sum_{i=1}^m k_i$ constraints.