ORF522 – Linear and Nonlinear Optimization

22. Robust Optimization

Bartolomeo Stellato — Fall 2024

Decision-making under uncertainty

so far we have modeled and solved optimization problems of the form

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x) \leq 0 \end{array}$$

what if we do not know f or g exactly?

Today's lecture

[RO, Ch 1 and 2] [TARO][ee364b]

Decision-making under uncertainty with Robust Optimization

- Example why uncertain data can be problematic
- Reformulating robust constraints
- Constructing uncertainty sets

Why is uncertain data an issue?

A medical production example

x_1 purchase raw material 1 (kg)

x_2 purchase raw material 2 (kg)

x_3 produce drug 1 (1000 packs)

x_4 produce drug 2 (1000 packs)

minimize	$c^T x$	
subject to	$-0.01x_1 - 0.02x_2 + 0.5x_3 + 0.6x_4 \leq 0$	balance of active agent
	$x_1 + x_2 \leq 1000$	storage
	$90x_3 + 100x_4 \leq 2000$	manpower
	$40x_3 + 50x_4 \leq 800$	equipment
	$100x_1 + 199.9x_2 + 700x_3 + 800x_4 \leq 800$	budget
	$x \geq 0$	

optimal objective -8819.66 (8.82% profit)

Uncertainty in the balance of active ingredients constraint

$$-0.01x_1 - 0.02x_2 + 0.5x_3 + 0.6x_4 \leq 0$$

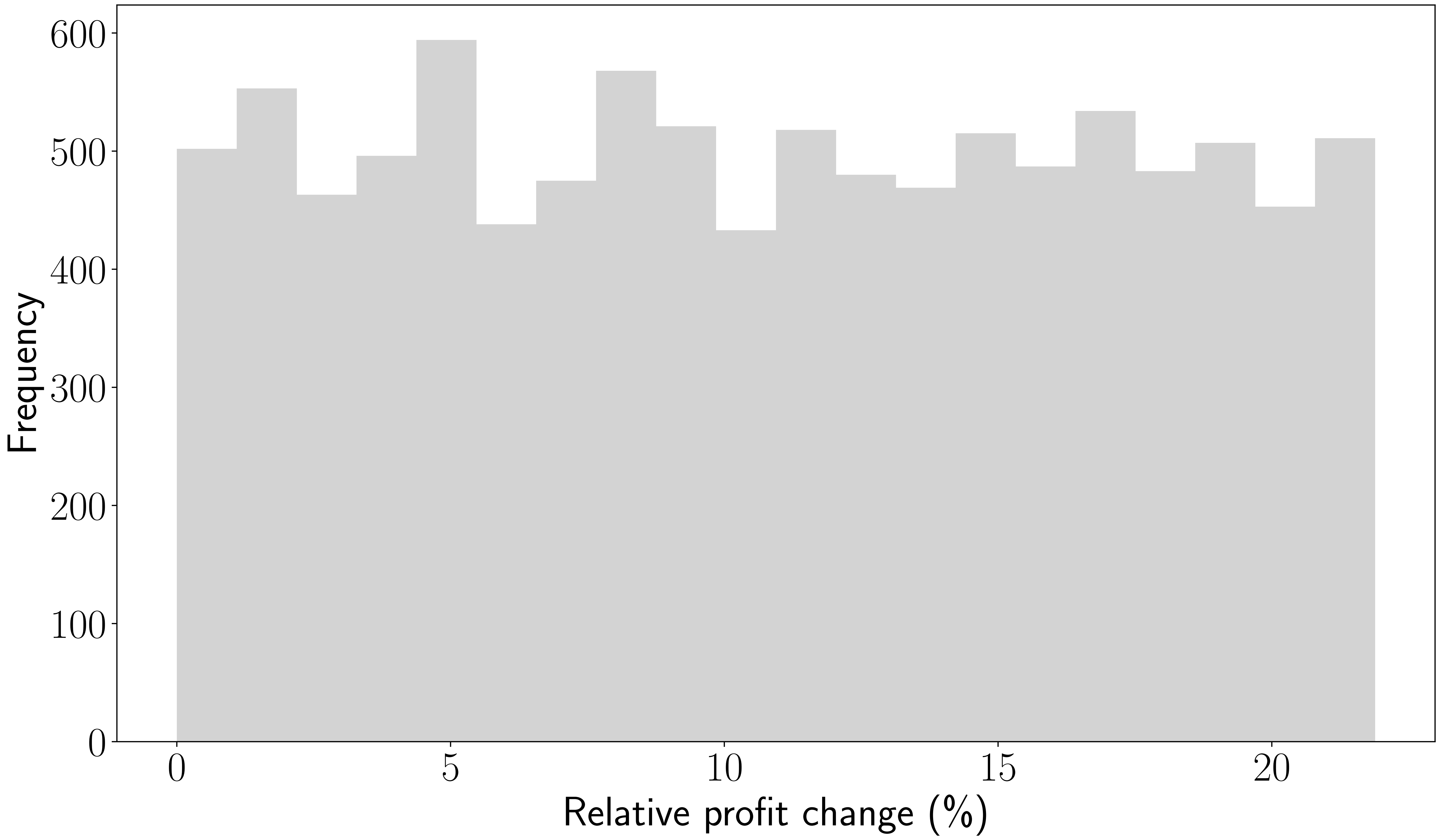
active agent content
in raw materials

What if the composition is not certain?

$-0.01(1 \pm 0.5\%)$ for raw material 1
$-0.02(1 \pm 2\%)$ for raw material 2

Constraint may become infeasible! → we adjust x_3 (prod. of drug 1) to ensure feasibility → What happens to the objective then?

Large suboptimality



Frequently lose up to more than 20% of profits

Chance constrained programs

Let $u \in \mathbf{R}^p$ be a random variable, we write a *chance constrained program* as

$$
\begin{array}{ll}
\text{minimize} & f(x) \\
\text{subject to} & \boxed{\mathbf{P}(g(u, x) \leq 0) \geq 1 - \epsilon}
\end{array}
$$

chance constraints

concave in first argument *(e.g., linear, bilinear)*

Remarks

- Can model uncertain objectives through epigraph forms
- Equality constraints do not make sense

Issues

- Typically intractable (nonconvex, except special cases)
- We often do not know the distribution

Robust Optimization

Replace *chance constraint* with

$$g(u, x) \le 0, \quad \forall u \in \boxed{\mathcal{U}}$$

uncertainty set

equivalently

$$\sup_{u \in \mathcal{U}} g(u, x) \le 0$$

Questions

- How to reformulate robust constraints? (when are they tractable?)
- How to construct the uncertainty sets?

Reformulating robust constraints

Reformulation of uncertain nonlinear constraints

Let $\mathcal{U}$ be a nonempty, convex, compact set satisfying mild regularity conditions. Then, $x \in \mathbf{R}^n$ satisfies $g(u, x) \leq 0,\ \forall u \in \mathcal{U}$ if and only if $\exists w$:

$$[-g]^*(-w, x) + \sigma_{\mathcal{U}}(w) \leq 0$$

conjugate function

$$[-g]^*(-w, x) = \sup_{z \in \mathbf{R}^n} -w^T z + g(z, x)$$

support function

$$\sigma_{\mathcal{U}}(w) = \sup_{u \in \mathcal{U}} w^T u = \mathcal{I}_{\mathcal{U}}^*(w)$$

Proof

$$\sup_{u \in \mathcal{U}} g(u, x) = \sup_{u} g(u, x) - \mathcal{I}_{\mathcal{U}}(u) = -\inf_{u} -g(u, x) + \mathcal{I}_{\mathcal{U}}(u)$$

$$= -\sup_{w} \left(-[-g]^*(-w, x) - \sigma_{\mathcal{U}}(w)\right) = \inf_{w} [-g]^*(-w, x) + \sigma_{\mathcal{U}}(w) \quad \blacksquare$$

Fenchel dual function
Lecture 15

Deriving robust counterparts of nonlinear uncertain inequalities, Ben-Tal, den Hertog, Vial

Constraint function examples

Bilinear functions: $g(u, x) = u^T P x + q^T u + r^T x + s$

$$[-g]^*(-w, x) = \sup_z -w^T z + z^T P x + q^T z + r^T x + s = \begin{cases} r^T x + s & w = Px + q \\ \infty & \text{otherwise} \end{cases}$$

Linear functions ($P = I, q = 0, r = a, s = -b$): $g(u, x) = (a + u)^T x - b$

$$[-g]^*(-w, x) = \begin{cases} a^T x - b & w = x \\ \infty & \text{otherwise} \end{cases}$$

Uncertainty set examples

Norm ball: $\mathcal{U} = \{u \in \mathbf{R}^p \mid \|u\| \leq \rho\}$

$$\sigma_{\mathcal{U}}(w) = \sup_{\|z\| \leq \rho} w^T z = \rho \sup_{\|v\| \leq 1} w^T v = \rho \underbrace{\|w\|_*}_{\text{dual norm}}$$

primal/dual norms
(Lecture 15)

Norm	Dual norm
2	2
1	∞
∞	1

Polyhedral set: $\mathcal{U} = \{u \in \mathbf{R}^p \mid Fu \leq g\}$

$$\sigma_{\mathcal{U}}(w) = \sup \{w^T z \mid Fz \leq g\} \underbrace{=}_{\text{strong duality LP}} \inf \{g^T \lambda \mid F^T \lambda = w,\ \lambda \geq 0\}$$

You can mix and match various functions and sets

Requirements

- $g(\cdot, x)$ concave and $[-g]^*(\cdot, x)$ easily computable for any x
- $\mathcal{U}$ is nonempty, convex, compact, and $\sigma_{\mathcal{U}}$ easily computable

You can break the rules sometimes and still get a convex problem (S-Lemma)

$$g(u, x) = \|Ax + Bu\|_2 \leq 1 \qquad\qquad \mathcal{U} = \{u \mid \|u\|_2 \leq 1\}$$

(*convex in* u)

In general, it is hard when $g(\cdot, x)$ not concave [More details in ORF523]

Back to production example

balance of active ingredients

$$-0.01x_1 - 0.02x_2 + 0.5x_3 + 0.6x_4 \le 0$$

uncertainty

raw material 1: $-0.01 \pm 5 \cdot 10^{-5}$

raw material 2: $-0.02 \pm 4 \cdot 10^{-4}$

$g(u, x)$ **equivalent robust constraint**

$$(a + Mu)^T x \le 0, \ \forall u \in \mathcal{U}$$

$$a = (-0.01, -0.02, 0.5, 0.6)$$

$$M = \mathbf{diag}(5 \cdot 10^{-5}, 4 \cdot 10^{-4}, 0, 0)$$

$$\mathcal{U} = \{u \in \mathbf{R}^4 \mid \|u\|_\infty \le 1\}$$

conjugate function

$$[-g]^*(-w, x) = \begin{cases} a^T x & w = M^T x \\ \infty & \text{otherwise} \end{cases}$$

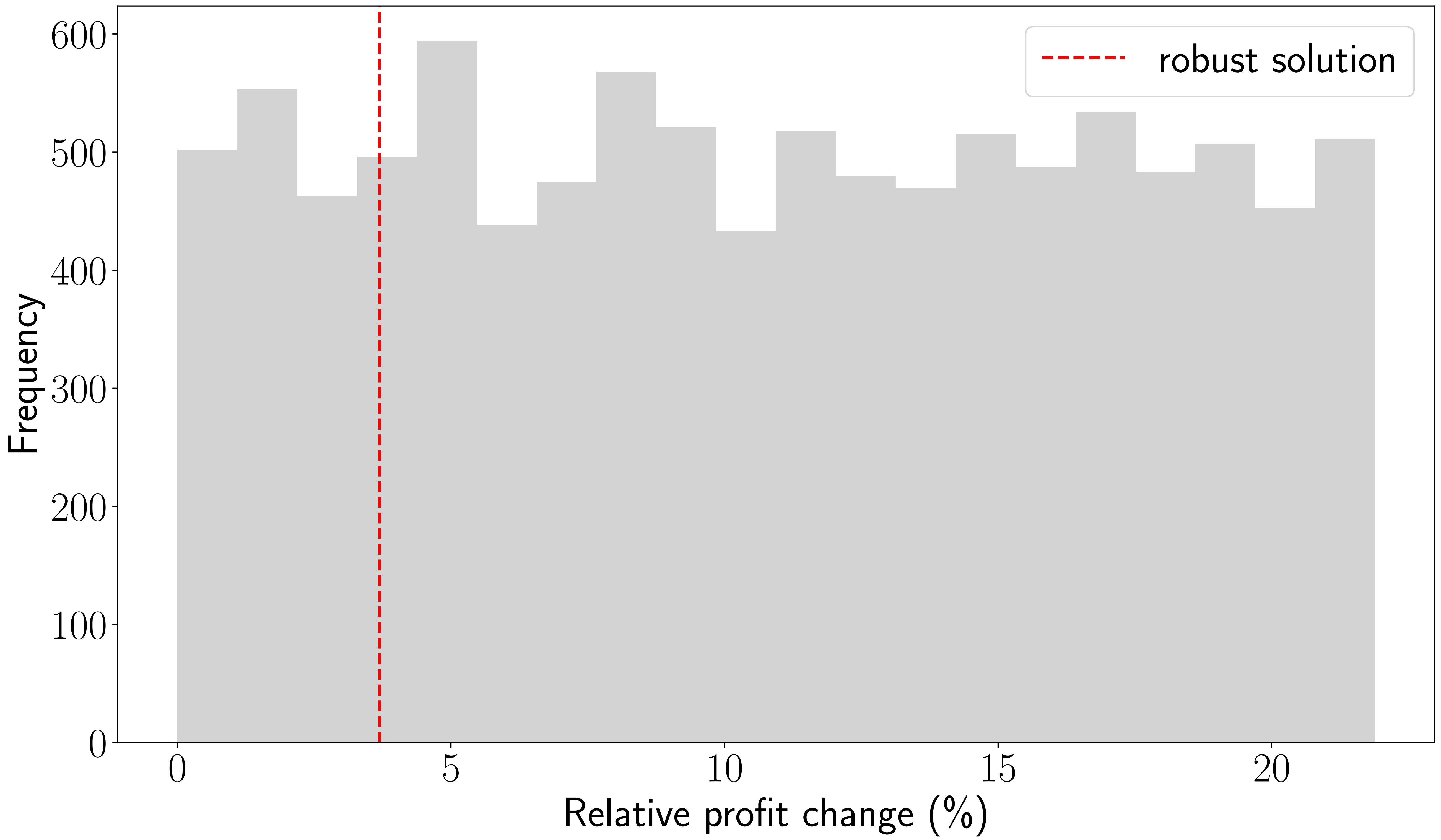
support function

$$\sigma_{\mathcal{U}}(w) = \|w\|_1$$

reformulation

$$a^T x + \|M^T x\|_1 \le 0$$

Robust performance in production example



$x_{\mathrm{ro}}^{\star}$ has degradation provably no worse than 3.7%

How can we construct $\mathcal{U}$ in general?

Constructing uncertainty sets

Uncertain linear constraint with Gaussian data

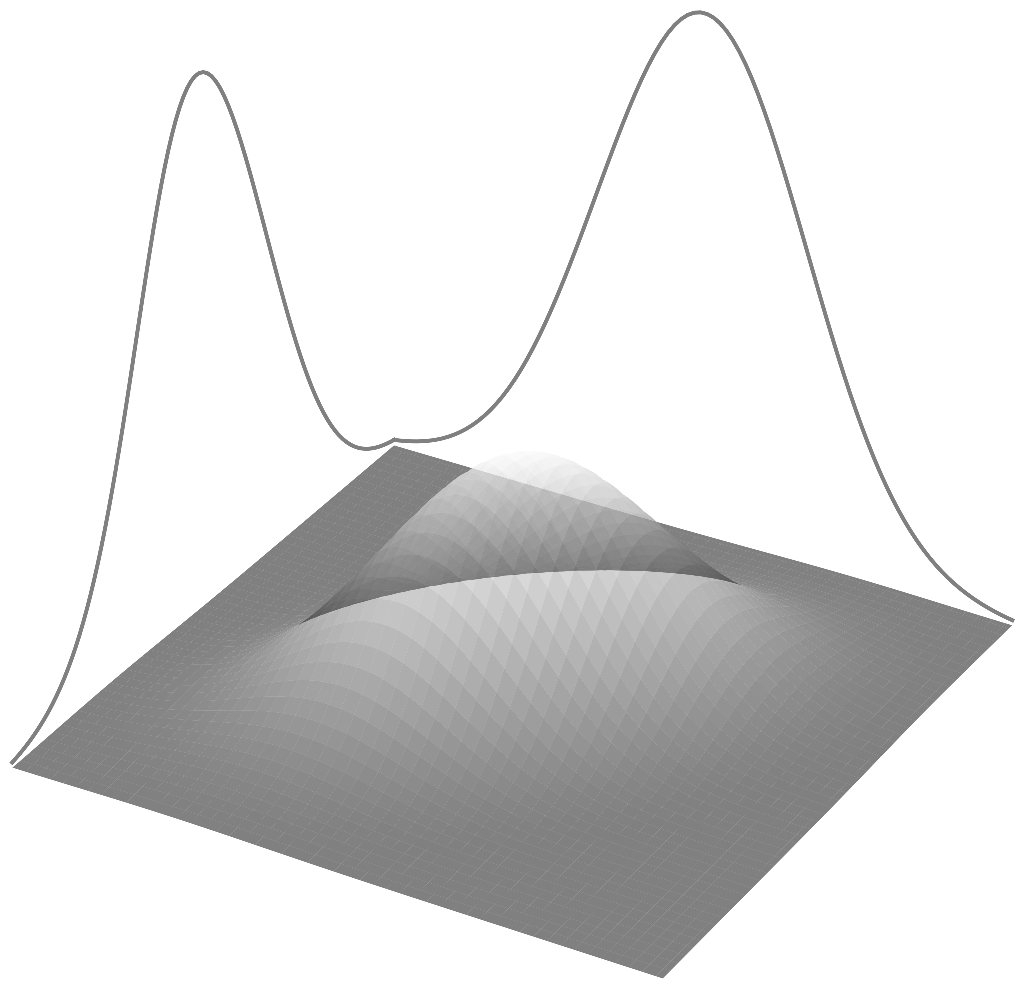
$$\mathbf{P}(a^T x \le b) \ge 1 - \epsilon$$

assumption
$1 - \epsilon \ge 1/2$

How can we rewrite
the chance constraint?

assumption
multivariate Gaussian

$$a \sim \mathcal{N}(\bar{a}, \Sigma)$$



projection property

$$a^T x \sim \mathcal{N}(\bar{a}^T x, x^T \Sigma x)$$

Uncertain linear constraint with Gaussian data

Exact reformulation

projection property

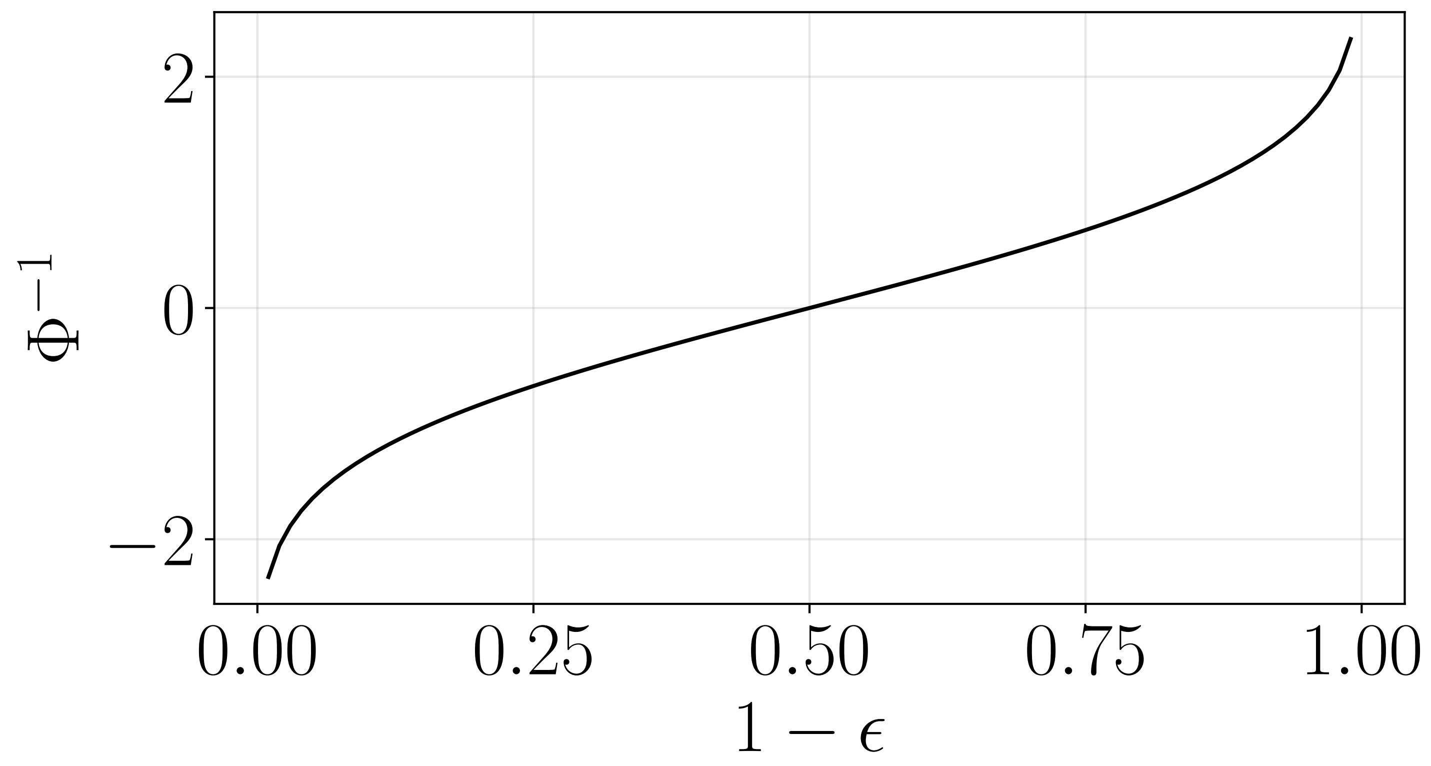
$$a^T x \sim \mathcal{N}(\bar{a}^T x, x^T \Sigma x)$$

$$\mathbf{P}(a^T x \leq b) = \mathbf{P}\left(\underbrace{\frac{a^T x - \bar{a}^T x}{\sqrt{x^T \Sigma x}}}_{\mathcal{N}(0,1)} \leq \frac{b - \bar{a}^T x}{\sqrt{x^T \Sigma x}}\right) = \Phi\left(\frac{b - \bar{a}^T x}{\sqrt{x^T \Sigma x}}\right)$$

Φ: cumulative distribution function of zero mean unit variance Gaussian

$$\Phi(z) = \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt$$

$$\mathbf{P}(a^T x \leq b) \geq 1 - \epsilon \quad \Longleftrightarrow \quad \frac{b - \bar{a}^T x}{\sqrt{x \Sigma x}} \geq \Phi^{-1}(1 - \epsilon)$$



exact reformulation

$$\bar{a}^T x + \underbrace{\Phi^{-1}(1 - \epsilon)}_{\geq 0} \|\Sigma^{1/2} x\|_2 \leq b$$

equivalent to robust constraint with

$$g(u, x) = (\bar{a} + \Sigma^{1/2} u)^T x - b$$

$$\mathcal{U} = \{u \in \mathbf{R}^n \mid \|u\|_2 \leq \Phi^{-1}(1 - \epsilon)\}$$

Approximation with high probability bounds

$$\mathbf{P}(a^T x \leq b) \geq 1 - \epsilon$$

What if the distribution is not Gaussian? $\longrightarrow$ Approximate chance constraint with high probability bounds

Assumption
mean $\bar{a}$, and $a \in [\ell, u]$ $\longrightarrow$

$$\mathbf{E}(a_i x_i) = \bar{a}_i x_i$$
$$\ell_i x_i \leq a_i x_i \leq u_i x_i$$

Hoeffding's inequality

$$\mathbf{P}(a^T x \leq b) = \mathbf{P}\left(\sum_{i=1}^n a_i x_i - \bar{a}_i x_i \leq b - \bar{a}^T x\right) \underset{\text{Hoeffding's inequality}}{\boxed{\geq}} 1 - \exp\left(-\frac{2(b - \bar{a}^T x)^2}{\sum_{i=1}^n (u_i - \ell_i)^2 x_i^2}\right)$$

more compactly: $\quad \mathbf{P}(a^T x \leq b) \geq 1 - \exp\left(-\frac{2(b - \bar{a}^T x)^2}{\|\mathbf{diag}(u - \ell)x\|_2^2}\right)$

Constraint reformulation with high probability bounds

$$\mathbf{P}(a^T x \le b) \ge 1 - \exp\left(-\frac{2(b - \bar{a}^T x)^2}{\|\mathbf{diag}(u-\ell)x\|_2^2}\right) \ge 1 - \epsilon$$

$$\iff \quad \exp\left(-\frac{2(b - \bar{a}^T x)^2}{\|\mathbf{diag}(u-\ell)x\|_2^2}\right) \le \epsilon$$

$$\iff \quad -\frac{2(b - \bar{a}^T x)^2}{\|\mathbf{diag}(u-\ell)x\|_2^2} \le \log(\epsilon)$$

$$\iff \quad 2(b - \bar{a}^T x)^2 \ge \log(1/\epsilon)\|\mathbf{diag}(u-\ell)x\|_2^2$$

reformulation

$$\bar{a}^T x + \sqrt{\frac{1}{2}\log\frac{1}{\epsilon}}\|\mathbf{diag}(u-\ell)x\|_2 \le b$$

equivalent to robust constraint with

$$g(v, x) = (\bar{a} + \mathbf{diag}(u-\ell)v)^T x - b$$

$$\mathcal{U} = \left\{u \in \mathbf{R}^n \mid \|u\|_2 \le \sqrt{\frac{1}{2}\log\frac{1}{\epsilon}}\right\}$$

Portfolio optimization example

asset allocations $x \in \mathbf{R}^n$
uncertain returns $r \in [\ell, u]$ with mean μ

optimization problem

$$
\begin{array}{ll}
\text{maximize} & \underbrace{\mu^T x}_{\text{expected returns}} \\
\text{subject to} & \underbrace{\mathbf{P}(r^T x \le \underbrace{\alpha}_{\text{unwanted return level}}) \le \epsilon}_{\text{loss risk constraint}} \\
& \mathbf{1}^T x = 1 \\
& x \ge 0
\end{array}
$$

robust reformulation

$$
\begin{array}{ll}
\text{maximize} & \mu^T x \\
\text{subject to} & \mu^T x - \sqrt{(1/2)\log(1/\epsilon)}\|\mathbf{diag}(u-\ell)x\|_2 \ge \alpha \\
& \mathbf{1}^T x = 1 \\
& x \ge 0
\end{array}
$$

Portfolio optimization strategies

returns

$$\mu_i = 1.05 + \frac{3(n-i)}{10n} \quad (\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n)$$

$$|r_i - \mu_i| \leq u_i = 0.05 + \frac{n-i}{2n}, \qquad u_n = 0$$

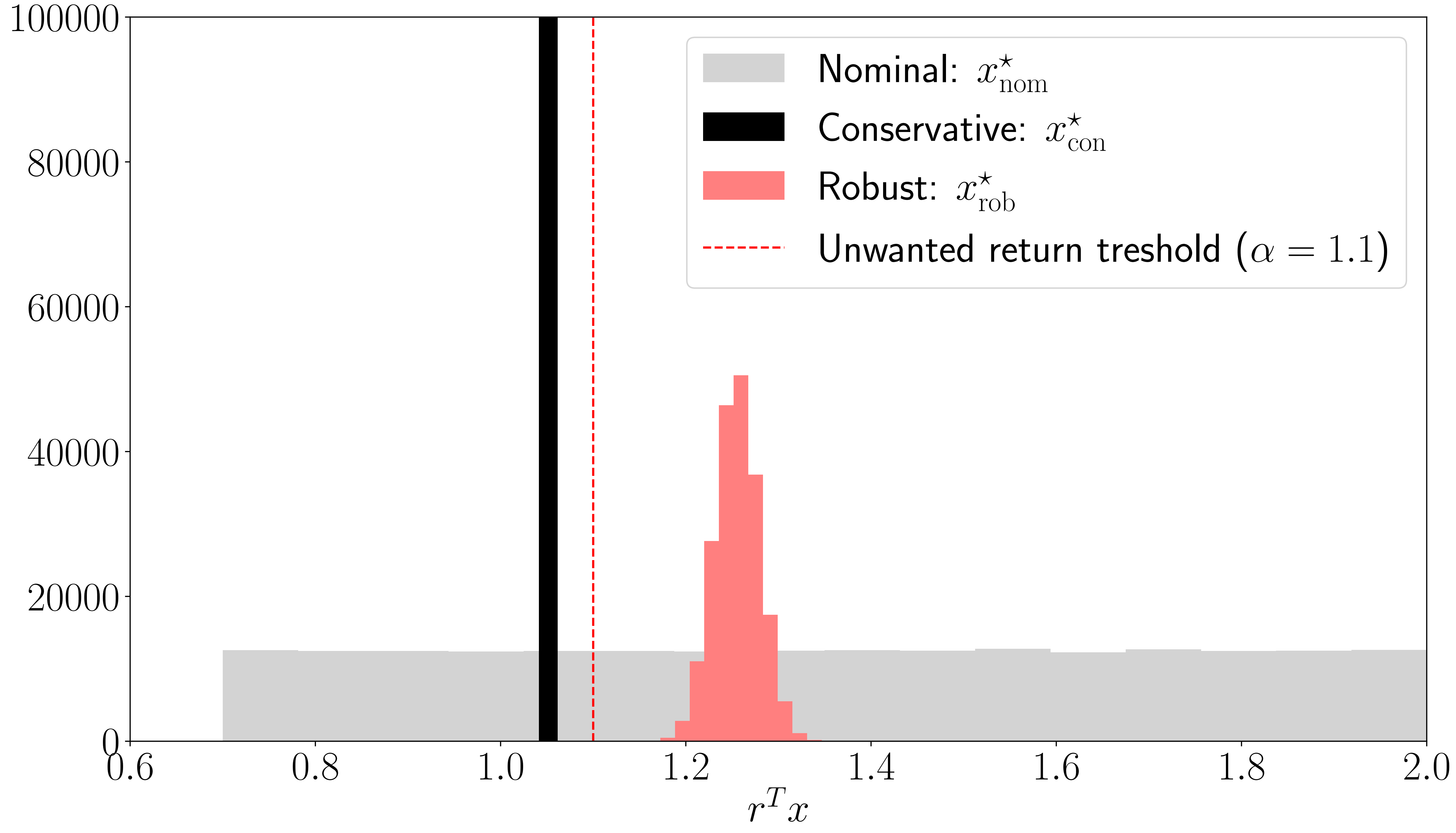
n-th asset is cash
(guaranteed 5% return)

baselines

- nominal minimizer $x^\star_{\text{nom}} = e_1$
- conservative minimizer $x^\star_{\text{con}} = e_n$
- robust minimizer $x^\star_{\text{ro}}$

Portfolio optimization results comparison

$$\mathbf{P}(r^T x \le \underset{1.1}{\alpha}) \le \underset{10^{-3}}{\epsilon}$$



robust is guaranteed to do better than conservative (cash)

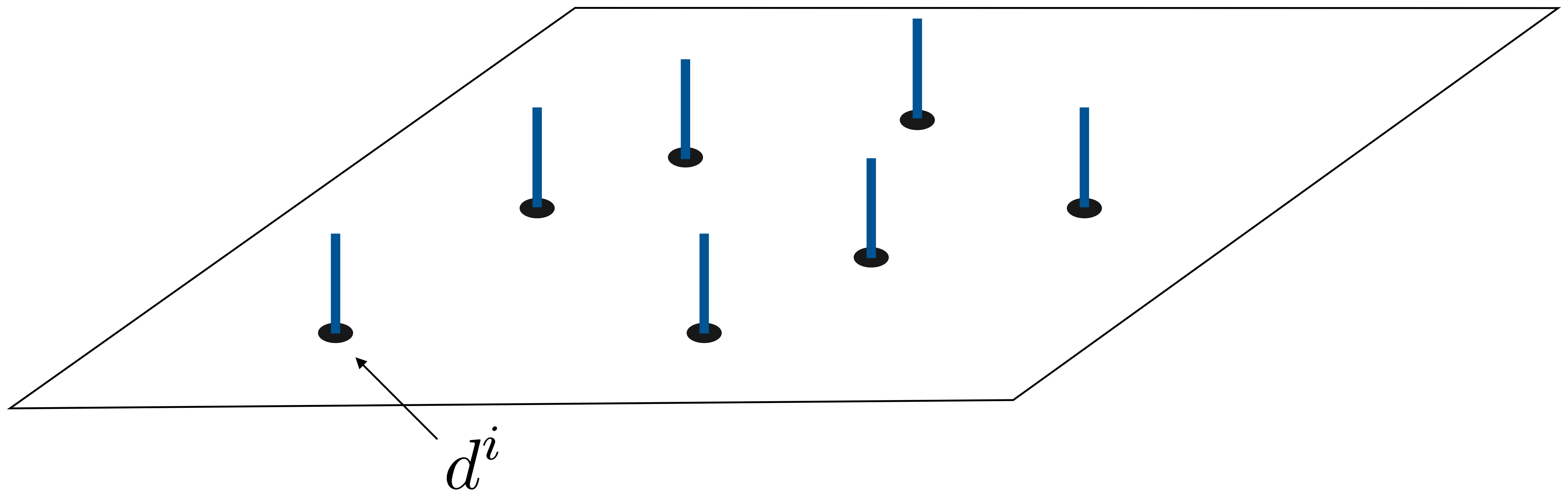
Issues with traditional uncertainty set construction

Probability distributions $\mathbf{P}$ are never observed in practice

Data is observed in practice

All we have is the empirical distribution

$$\hat{\mathbf{P}}^N = \frac{1}{N}\sum_{i=1}^{N}\delta_{d^i}$$



Are there better ways to model the uncertainty
that still lead to tractable formulations?

Robust Optimization

- Today we learned to:
  - **Understand** the limitations of optimization in presence of uncertainty
  - **Derive** tractable reformulations of robust constraints using duality theory
  - **Construct** uncertainty sets from probabilistic assumptions on the uncertainty

Next lecture

- Bringing data in the picture