ECE 602 – Section 5
Standard optimization problems

- Standard optimization problem
- Equivalent optimization problems
- Linear and Quadratic Programming
- Conic programming and SOCP
- Semidefinite Programming

- *Standard optimization problem* has the following form

$$\begin{aligned} &\min_x \; f_0(x) \\ &\text{subject to } f_i(x) \le 0, \quad i = 1, \ldots, m \\ &\qquad\qquad\quad h_j(x) = 0, \quad j = 1, \ldots, p \end{aligned}$$

where:

- $x \in \mathbf{R}$ is the *optimization variable*
- $f_0 : \mathbf{R}^n \to \mathbf{R}$ is the *cost* (or *objective*) *function*
- $f_i : \mathbf{R}^n \to \mathbf{R}$ are *inequality constraint functions*
- $h_j : \mathbf{R}^n \to \mathbf{R}$ are *equality constraint functions*

- The *domain* of the problem is $\mathcal{D} = \bigcap_{i=0}^{m} \mathbf{dom}\, f_i \cap \bigcap_{j=1}^{p} \mathbf{dom}\, h_j$.

- The *optimal value* is

$$\boxed{p^* = \inf_x \{f_0(x) \mid f_i(x) \le 0, h_j(x) = 0, \forall i, j\}}$$

- A point $x \in \mathcal{D}$ is *feasible* if it satisfies all the constraints.
- The problem is *feasible* if the set of feasible points is non-empty. If the problem is infeasible, we have $p^* = \infty$.
- x^* is an *optimal point*, if x^* is *feasible* and $f_0(x^*) = p^*$.
- The set of all optimal points is the *optimal set*

$$X_{opt} = \{x \mid x \text{ is feasible and } f(x) = p^*\}$$

- A feasible x is locally optimal, if $\exists\, R > 0$ such that x is optimal for

$$\begin{aligned} \min_z \;& f_0(z) \\ \text{subject to } & f_i(z) \leq 0, \quad i = 1, \ldots, m \\ & h_j(z) = 0, \quad j = 1, \ldots, p \\ & \boxed{\|z - x\|_2 \leq R} \end{aligned}$$

- This means x minimizes f_0 over nearby points in the feasible set.

Some examples

Consider the following *unconstrained* scalar problems.

- $f_0(x) = 1/x$ with $\mathbf{dom}\, f_0 = \mathbf{R}_+$

  In this case, $p^* = 0$, but there is no optimal point.

- $f_0(x) = -\log x$ with $\mathbf{dom}\, f_0 = \mathbf{R}_+$

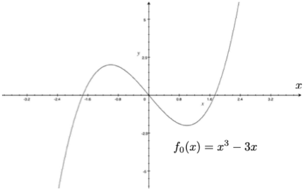
  In this case, the cost function is unbounded below and thus $p^* = -\infty$.

- $f_0(x) = x \log x$ with $\mathbf{dom}\, f_0 = \mathbf{R}_+$

  This problem has an optimal $p^* = -1/e$ attained at *unique* $x^* = 1/e$.

- $f_0(x) = x^3 - 3x$ with $\mathbf{dom}\, f_0 = \mathbf{R}$

  This problem is unbounded from below, i.e. $p^* = -\infty$. However, there is a local minimum at $x^*_{\text{loc}} = 1$.



- If the objective function is identically zero (i.e., $f_0 \equiv 0$), then

$$p^* = \begin{cases} 0, & \text{feasible set} \neq \emptyset \\ \infty, & \text{feasible set} = \emptyset \end{cases}$$

- Thus, the standard-form optimization problem

$$\begin{aligned} &\min_x \; 0 \\ &\text{subject to } f_i(x) \le 0, \quad i = 1, \ldots, m \\ &\qquad\qquad\quad h_j(x) = 0, \quad j = 1, \ldots, p \end{aligned}$$

is equivalent to the *feasibility problem* that is given by

$$\begin{aligned} &\text{find } x \\ &\text{subject to } f_i(x) \le 0, \quad i = 1, \ldots, m \\ &\qquad\qquad\quad h_j(x) = 0, \quad j = 1, \ldots, p \end{aligned}$$

which aims at finding a *feasible* (not optimal!) solution.

- Consider the following example with *box constraints:*

$$\min_x \; f_0(x)$$
$$\text{subject to } l_i \le x_i \le u_i, \quad i = 1, \dots, n$$

- We can express this problem in a standard form as

$$\begin{aligned} &\min_x \; f_0(x) \\ &\text{subject to } l_i - x_i \le 0, \quad i = 1, \dots, n \\ &\qquad\qquad\quad x_i - u_i \le 0, \quad i = 1, \dots, n \end{aligned}$$

- In this case, we have $2n$ equality constraint functions of the form

$$\begin{aligned} f_i(x) &= l_i - x_i, \quad i = 1, \dots, n \\ f_i(x) &= x_i - u_i, \quad i = n+1, \dots, 2n \end{aligned}$$

- Suppose $\phi : \mathbf{R}^n \to \mathbf{R}^n$ is one-to-one with $\phi(\mathbf{dom}\,\phi) \subseteq \mathcal{D}$.
- Define

$$\tilde{f}_i(z) = f_i(\phi(z)), \ \forall i \quad \text{and} \quad \tilde{h}_j(z) = h_j(\phi(z)), \ \forall j$$

- Then the problem

$$\begin{aligned} \min_z \ & \tilde{f}_0(z) \\ \text{subject to } & \tilde{f}_i(z) \leq 0, \quad i = 1, \ldots, m \\ & \tilde{h}_j(z) = 0, \quad j = 1, \ldots, p \end{aligned}$$

  is related to the original problem by the *change of variables* $x = \phi(z)$.

- If z solves the above problem, then $x = \phi(z)$ solves the original one.
- If x solves the original problem, then $z = \phi^{-1}(x)$ solves the new one.

- Suppose that $\psi_0 : \mathbf{R} \to \mathbf{R}$ is monotone increasing, and
  - $\psi_1, \ldots, \psi_m : \mathbf{R} \to \mathbf{R}$ satisfy $\psi_i(u) \leq 0$ iff $u \leq 0$
  - $\psi_{m+1}, \ldots, \psi_{m+p} : \mathbf{R} \to \mathbf{R}$ satisfy $\psi_i(u) = 0$ iff $u = 0$
- Define the following functions

$$\tilde{f}_i(x) = \psi_i(f_i(x)), \quad i = 0, \ldots, m$$
$$\tilde{h}_j(x) = \psi_{m+i}(h_i(x)), \quad j = 0, \ldots, p$$

- Then the problem

$$\begin{aligned} \min_x \ & \tilde{f}_0(x) \\ \text{subject to } & \tilde{f}_i(x) \leq 0, \quad i = 1, \ldots, m \\ & \tilde{h}_j(x) = 0, \quad j = 1, \ldots, p \end{aligned}$$

  is equivalent to the original problem.

- For example, $\min_x \|Ax - b\|_2$ and $\min_x \|Ax - b\|_2^2$ are equivalent.

- As a simple example, consider the unconstrained problem

$$\min_x f(x), \quad \text{where } f(x) = \begin{cases} x^T x, & Ax = b \\ \infty, & \text{otherwise} \end{cases}$$

- The problem has an *implicit equality constraint* $Ax = b$.

- The implicit constraint can be made *explicit* by considering instead

$$\min_x x^T x$$
$$\text{subject to } Ax = b$$

- The problems are clearly equivalent, they are not the same.

- While the first problem is unconstrained and non-differentiable, the second is constrained and differentiable.

- A *convex optimization problem* is one of the form

$$\begin{aligned} \min_x \; & f_0(x) \\ \text{subject to } & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & a_j^T x = b_j, \quad j = 1, \ldots, p \end{aligned}$$

where $f_0, f_1, \ldots, f_m$ are ȝconvex functions.

- The feasible set of a convex optimization problem is *convex*.
- Thus, in a convex optimization problem, *we minimize a convex cost function over a convex set*.
- If $f_0(x)$ is concave, we replace minimization by maximization in order to still regard the problem as "convex".

- Consider the following problem

$$\min_x \; f_0(x) = x_1^2 + x_2^2$$
$$\text{subject to } f_1(x) = x_1/(1 + x_2^2) \leq 0$$
$$h_1(x) = (x_1 + x_2)^2 = 0$$

- This problem is *not* a convex optimization problem in standard form, although the feasible set $\{x \mid x_1 \leq 0, \, x_1 + x_2 = 0\}$ is convex.

- The problem can be reformulated as

$$\min_x \; f_0(x) = x_1^2 + x_2^2$$
$$\text{subject to } f_1(x) = x_1 \leq 0$$
$$h_1(x) = x_1 + x_2 = 0$$

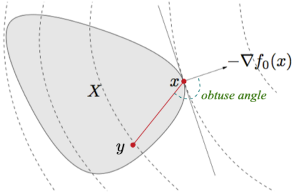
which is a convex optimization problem in standard form.

- A *fundamental property* of convex optimization problems is that *their locally optimal points are also globally optimal*.
- Suppose that f_0 is differentiable, so that for all $x, y \in \mathbf{dom}\, f_0$,

$$f_0(y) \geq f_0(x) + \nabla f_0(x)^T (y - x)$$

- Let $\mathcal{X}$ denote the feasible set. Then, x is *optimal* if and only if $x \in \mathcal{X}$ and

$$-\nabla f_0(x)^T (y - x) \leq 0, \quad \forall y \in \mathcal{X}$$



- Thus, if nonzero, $\nabla f_0(x)$ *defines a supporting hyperplane* to $\mathcal{X}$ at x.
- Note that, when $\mathcal{X} = \mathbf{R}^n$ (i.e. no constraints), the above condition is trivially reduced to $\nabla f_0(x) = \bar{0}$.

- **Unconstrained problem:** $\min f_0(x)$

  In this case, x is optimal if and only if

  $$x \in \mathbf{dom}\, f_0 \quad \& \quad \nabla f_0(x) = 0$$

- **Equality constrained problem:** $\min f_0(x)$ subject to $Ax = b$

  Here, x is optimal if and only if $\exists\, \nu$ such that

  $$x \in \mathbf{dom}\, f_0 \quad \& \quad Ax = b \quad \& \quad \nabla f_0(x) + A^T \nu = 0$$

  Note that the latter implies that $-\nabla f_0(x) \in \mathrm{Col}(A^T)$ $(= \mathrm{Row}(A))$.

- **Minimization over $\mathbf{R}_+$:** $\min f_0(x)$ subject to $x \succeq 0$

  In this case, x is optimal if and only if

  $$x \in \mathbf{dom}\, f_0 \quad \& \quad x \succeq 0 \quad \& \quad \begin{cases} \nabla f_0(x)_i \geq 0, & x_i = 0 \\ \nabla f_0(x)_i = 0, & x_i > 0 \end{cases}$$

  (For more details, see Section 4.2.3 of Boyd's textbook.)

- Consider the following optimization problem:

$$\min_{x} \; f_0(x)$$
$$\text{subject to } Ax = b$$

where $A \in \mathbf{R}^{m\times n}$, with $\mathrm{rank}(A) = m < n$.

- Since A is "wide", there are infinitely many solutions to $Ax = b$. Let x_0 be a *particular solution* to this system (e.g., $x_0 = A^T(AA^T)^{-1}b = A^\dagger b$). Then, a *general solution* to $A\,x = b$ can be expressed as $x = x_0 + x_{\mathrm{hom}}$, where x_{hom} is *any* solution to $A\,x = 0$.

- Recall that the space of the *homogeneous solutions* $\mathcal{N}_A = \{u \mid Au = 0\}$ is a linear subspace of $\mathbf{R}^n$ of dimension $p = n - m$. Hence, any $u \in \mathcal{N}_A$ can be represented as $u = Bz$, with the columns of $B \in \mathbf{R}^{n\times p}$ forming a basis in $\mathcal{N}_A$.

- Then, solutions to the above problem can be expressed as $x = x_0 + Bz$, thus reducing it to

$$\min_{z} \; f_0(Bz + x_0)$$

with $x^* = Bz^* + x_0$.

- **Introducing equality constraints:**

$$\min_{x} \; f_0(A_0 x + b_0)$$
$$\text{subject to } f_i(A_i x + b_i) \leq 0, \quad i = 1, \ldots, m$$

is equivalent to

$$\min_{x,y} \; f_0(y_0)$$
$$\text{subject to } f_i(y_i) \leq 0, \quad i = 1, \ldots, m$$
$$y_j = A_j x + b_j, \quad j = 0, 1, \ldots, m$$

- **Introducing slack variables:**

$$\min_{x} \; f_0(x)$$
$$\text{subject to } a_i^T x \leq b_i, \quad i = 1, \ldots, m$$

is equivalent to

$$\min_{x,s} \; f_0(x)$$
$$\text{subject to } a_j^T x + s_j = b_j, \quad j = 1, \ldots, m$$
$$s_i \geq 0, \quad i = 1, \ldots, m$$

- **Epigraph form:**

$$\begin{aligned} &\min_{x,t} \; t \\ &\text{subject to } f_0(x) - t \leq 0 \\ &\qquad\qquad f_i(x) \leq 0, \quad i = 1, \ldots, m \\ &\qquad\qquad a_i^T x = b_i, \quad i = 1, \ldots, p \end{aligned}$$

- **Marginization:**

$$\begin{aligned} &\min_{x_1, x_2} \; f_0(x_1, x_2) \\ &\text{subject to } f_i(x_1) \leq 0, \quad i = 1, \ldots, p \end{aligned}$$

is equivalent to

$$\begin{aligned} &\min_{x_1} \; \tilde{f}_0(x_1) \\ &\text{subject to } f_i(x_1) \leq 0, \quad i = 1, \ldots, p \end{aligned}$$

where $\tilde{f}_0(x_1) = \inf_{x_2} f_0(x_1, x_2)$ (which is *unconstrained*).

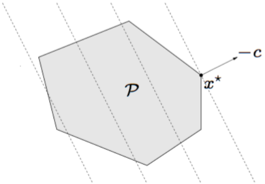
- A general *linear program* (LP) has the form

$$\min_x \; c^T x \quad \text{s.t.} \quad Ax \preceq b, \; Cx = d$$

where $A \in \mathbf{R}^{m \times n}$, $C \in \mathbf{R}^{p \times n}$, $b \in \mathbf{R}^m$ and $d \in \mathbf{R}^p$.

- A maximization problem with affine objective and constraint functions is an LP as well.

- The feasible set of an LP is a *polyhedron*.
- Numerically, optimal solutions occur at the *vertices*.



- If the LP has no equality constraints, it is called an *inequality form LP*

$$\min_x \; c^T x, \quad \text{s.t.} \quad Ax \preceq b$$

- Since $x = x^+ - x^-$, with $x^+ = \max\{x, 0\}$ and $x^- = \min\{x, 0\}$, the cost can be redefined as $c^T x = [c, -c]\,[x^+, x^-]^T$, with $[x^+, x^-]^T \succeq 0$.
- Let $s \succeq 0$ be a vector of slack variables such that $Ax + s = b$ and define a new optimization variable $y = [x^+, x^-, \, s]^T \succeq 0$.
- Then, with $C = [A, -A, \, I]$ and $d = [c, -c, \, 0]^T$, the original LP can be equivalently written as

$$\min_y \; d^T y, \quad \text{s.t.} \quad Cy = b, \;\; y \succeq 0$$

which is known as a *standard form LP*.

- **ℓ_∞-norm approximation**: $f(x) = \|A\,x - b\|_\infty$

  Let $\mathbf{1} \in \mathbf{R}^m$ be vector of ones. Then, the problem can be expressed as

$$\begin{aligned} \min_{x,t} \; & t \\ \text{s.t.} \; & Ax - b \preceq t\mathbf{1} \\ & Ax - b \succeq -t\mathbf{1} \end{aligned}$$

  Indeed, the constraints require that $|a_k^T x - b_k| \le t$, for all k, and thus

$$t \ge \max_k |a_k^T x - b_k| = \|Ax - b\|_\infty$$

- **ℓ_1-norm approximation**: $f(x) = \|A\,x - b\|_1$

  In this case, the equivalent LP is

$$\begin{aligned} \min_{x,s} \; & \mathbf{1}^T s \\ \text{s.t.} \; & Ax - b \preceq s \\ & Ax - b \succeq -s \end{aligned}$$

  where the constraints require that $|a_k^T x - b_k| \le s_k$, for all k.

- Let $\mathcal{C}$ be a *polyhedron* defined as

$$\mathcal{C} = \left\{ x \in \mathbf{R}^n \mid a_i^T x \le b_i,\ i = 1, 2, \ldots, m \right\}$$

- Consider the problem of finding the *largest ball* $\mathcal{B} = \{x_c + u \mid \|u\| \le r\}$ (with unknowns x_c and r) that lies inside of $\mathcal{C}$. In this case, x_c is referred to as the *Chebyshev centre* of $\mathcal{C}$.

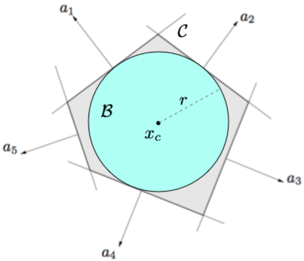
- Note that the inclusion requires that

$$a_i^T (x_c + u) \le b_i,\ i = 1, 2, \ldots, m$$

holds for any u. Therefore,

$$\sup_{u:\|u\| \le r} \left(a_i^T (x_c + u) \right) = a_i^T x_c + \overbrace{\sup_{u:\|u\| \le r} \left(a_i^T u \right)}^{r\|a_i\|_*}$$

thus requiring $a_i^T x_c + r\|a_i\|_* \le b_i$, $\forall i$.



- As a result, we arrive at an LP of the form

$$\max_{x_c, r} r \quad \text{subject to} \quad \begin{bmatrix} a_i \\ \|a_i\|_* \end{bmatrix}^T \begin{bmatrix} x_c \\ r \end{bmatrix} \le b_i,\ i = 1, 2, \ldots, m$$

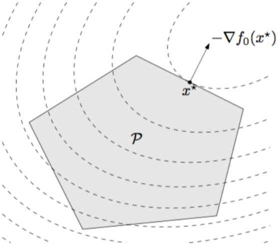
- A *quadratic program* (QP) is defined as

$$\min_x \frac{1}{2} x^T P x + q^T x + r$$
$$\text{s.t. } Ax \preceq b, \quad Cx = d$$

where $P \in \mathbf{S}^n_{++}$, $A \in \mathbf{R}^{m \times n}$ and $C \in \mathbf{R}^{p \times n}$.

- A *quadratically constrained quadratic program* (QCQP) has the form

$$\min_x \frac{1}{2} x^T P_0 x + q_0^T x + r_0$$
$$\text{s.t. } \frac{1}{2} x^T P_i x + q_i^T x + r_i \leq 0$$
$$Ax = b$$



where $P_i \in \mathbf{S}^n_{++}$, for all $i = 0, 1, \ldots, m$.

Example: least-squares and distance

- The problem of minimizing the convex quadratic function

$$\|Ax - b\|_2^2 = x^T A^T Ax - 2b^T Ax + b^T b$$

  is an (unconstrained) QP.

- The problem has many names, e.g., *regression analysis* or *least-squares (LS) approximation*. It is solved by $x = A^\dagger b$.

- *Constrained regression* or *constrained least-squares* is defined as

$$\begin{aligned} \min_x \quad & \|Ax - b\|_2^2 \\ \text{s.t.} \quad & l_i \le x_i \le u_i, \quad i = 1, \ldots, n \end{aligned}$$

  which is also a QP.

- Let $\mathcal{C}_1 = \{x \mid A_1 x \preceq b_1\}$ and $\mathcal{C}_2 = \{y \mid A_2 y \preceq b_2\}$ be two polyhedra. To find the *distance* between $\mathcal{C}_1$ and $\mathcal{C}_2$, one needs to solve

$$\begin{aligned} \min_{x,y} \quad & \|x - y\|_2^2 \\ \text{s.t.} \quad & A_1 x \preceq b_1, \; A_2 y \preceq b_2 \end{aligned}$$

  which is a QP.

- Consider a general LP

$$\min_x \; c^T x$$
$$\text{s.t. } Gx \preceq h$$
$$Ax = b$$

- Suppose $c \in \mathbf{R}^n$ is *random*, with mean value $\mathbf{E}\{c\} = \bar{c}$ and variance $\mathbf{E}\{(c - \bar{c})(c - \bar{c})^T\} = \Sigma \succ 0$.
- Note that $\mathbf{E}\{c^T x\} = \bar{c}^T x$ and $\mathbf{var}(c^T x) = x^T \Sigma x$.
- Define the *risk sensitive cost* as $\mathbf{E} c^T x + \gamma \, \mathbf{var}(c^T x)$, where $\gamma > 0$ controls the trade-off between expected cost and variance (risk).
- To minimize the risk-sensitive cost, we solve the following QP:

$$\min_x \; \bar{c}^T x + \gamma \, x^T \Sigma \, x$$
$$\text{s.t. } Gx \preceq h$$
$$Ax = b$$

- The *second-order cone program* (SOCP) is defined as

$$\begin{aligned} \min_x \; & a_0^T x \\ \text{s.t.} \; & A_0\, x = b_0 \\ & \|A_i\, x + b_i\|_2 \le c_i^T x + d_i, \quad i = 1, \ldots, m \end{aligned}$$

  where $A_i \in \mathbf{R}^{n_i \times n}$ and $A_0 \in \mathbf{R}^{p \times n}$.

- Note that the constraint $\|A_i x + b_i\|_2 \le c_i^T x + d_i$ suggests that

$$\boxed{(A_i x + b_i, c_i^T x + d_i) \; \in \; \mathcal{C} = \{(x,t) \in \mathbf{R}^n \times \mathbf{R}_+ \mid \|x\|_2 \le t\}}$$

  For this reason, it is called a *second-order cone constraint*.

- SOCPs are more general than QCQPs (and, of course, LPs), *viz.*
  - SOCP yields a general LP, if $A_i = 0$, $i = 1, \ldots, m$
  - SOCP yields QCQP, if $c_i = 0$, $i = 1, \ldots, m$

- Consider an LP in inequality form

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & a_i^T x \leq b_i, \quad i = 1, 2, \ldots, m \end{aligned}$$

- Thus, for example, a_i and b_i could represent *system* parameters, with x being *design* variables required to minimize a linear cost.

- Let us assume that there is some *uncertainly* in the parameters a_i that is formally described by the requirement

$$a_i \in \mathcal{C}_i = \left\{ \bar{a}_i + P_i u \mid \|u\| \leq 1 \right\}$$

with some $P_i \in \mathbf{R}^{n \times n}$. (Note that $P_i = 0$ means a_i is known perfectly.)

- Consider the following *robust linear program*

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & a_i^T x \leq b_i \text{ for all } a_i \in \mathcal{C}_i, \; i = 1, 2, \ldots, m \end{aligned}$$

- The *robust linear constraint*, $a_i^T x \leq b_i$, $\forall a_i \in \mathcal{C}_i$, can be expressed as

$$\sup_{a_i \in \mathcal{C}_i} a_i^T x \leq b_i$$

or, more specifically,

$$\sup_{a_i \in \mathcal{C}_i} a_i^T x = \bar{a}_i^T x + \sup_{u:\|u\| \leq 1} (u^T P^T x) = \bar{a}_i^T x + \|P_i^T x\|_* \leq b_i$$

which is a *second-order cone constraint* when $\|\cdot\| = \|\cdot\|_2 = \|\cdot\|_*$.

- In this case, the robust LP can be expressed as the SOCP

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & \bar{a}_i^T x + \|P_i^T x\|_2 \leq b_i, \; i = 1, 2, \ldots, m \end{aligned}$$

- The terms $\|P_i^T x\|_2$ act as *regularization terms*, preventing x from being large in directions with considerable uncertainty in a_i.

- Consider the case when a_i are *independent Gaussian random vectors*, with mean $\bar{a}_i \in \mathbf{R}^n$ and covariance $\Sigma_i \succ 0$, $a_i \sim \mathcal{N}(\bar{a}_i, \Sigma_i)$.

- We require that each constraint $a_i^T x \leq b_i$ holds with a probability (or *confidence*) exceeding $\eta \geq 0.5$, *viz.* $\mathbf{prob}\,(a_i^T x \leq b_i) \geq \eta$.

- This results in an *LP with random constraints*

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & \mathbf{prob}\,(a_i^T x \leq b_i) \geq \eta, \;\; i = 1, 2, \ldots, m \end{aligned}$$

- For any i, let $u = a_i^T x$. Note that u is a Gaussian random variable, with mean $\bar{u} = \bar{a}_i^T x$ and variance $\sigma^2 = x^T \Sigma_i x = \|\Sigma^{1/2} x\|_2^2$.

- Using u, the constraint $\mathbf{prob}\,(a_i^T x \leq b_i) \geq \eta$ can be expressed as

$$\mathbf{prob}\,(u \leq b_i) \geq \eta \quad \Longleftrightarrow \quad \mathbf{prob}\left(\underbrace{\frac{u - \bar{u}}{\sigma}}_{z} \leq \underbrace{\frac{b_i - \bar{u}}{\sigma}}_{z_i} \right) \geq \eta$$

- Define $z = (u - \bar{u})/\sigma$ and $z_i = (b_i - \bar{u})/\sigma$.
- Since $u \sim \mathcal{N}(\bar{u}, \sigma^2)$, then $z \sim \mathcal{N}(0, 1)$ and thus $\mathbf{prob}(z \leq z_i) = \Phi(z_i)$, where

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

  is the *cumulative distribution function* of the *normal* distribution.
- Since Φ is *strictly* monotone, the **prob** constraint can be expressed as

$$\Phi\left(\frac{b_i - \bar{u}}{\sigma}\right) \geq \eta \iff \frac{b_i - \bar{u}}{\sigma} \geq \Phi^{-1}(\eta) \iff \bar{u} + \Phi^{-1}(\eta)\sigma \leq b_i$$

  with $\Phi^{-1}(\eta) \geq 0$ (as $\eta > 0.5$).
- Hence, the LP with random constraints can be expressed as the SOCP

$$\begin{aligned} \min_{x} \quad & c^T x \\ \text{s.t.} \quad & \bar{a}_i^T x + \Phi^{-1}(\eta) \|\Sigma_i^{1/2} x\|_2 \leq b_i, \; i = 1, 2, \dots, m \end{aligned}$$

- A *semidefinite program* (SDP) has the form

$$\begin{aligned} \min_{x} \quad & c^T x \\ \text{s.t.} \quad & x_1 F_1 + x_2 F_2 + \dots x_n F_n + G \preceq 0 \\ & A\,x = b \end{aligned}$$

  where $G, F_1, F_2, \dots, F_n \in \mathbf{S}^m$ and $A \in \mathbf{R}^{p\times n}$. The above inequality is a *linear matrix inequality*.

- If the matrices $G, F_1, F_2, \dots, F_n$ are all diagonal, then the SDP reduces to a linear program, *viz.*

$$\begin{aligned} \min_{x} \quad & c^T x \\ \text{s. t.} \quad & x_1 F_1 + x_2 F_2 + \dots x_n F_n + G \preceq 0 \\ & Ax = b \end{aligned}$$

- A *standard form SDP* is defined as

$$\begin{aligned} \min_{X} \quad & \mathbf{tr}(CX) \\ \text{s.t.} \quad & \mathbf{tr}(A_i X) = b_i, \quad i = 1, \ldots, m \\ & X \succeq 0 \end{aligned}$$

where $C, A_1, \ldots, A_m \in \mathbf{S}^n$. This form is analogous to LP in standard (equality) form.

- An *inequality form SDP* is defined as

$$\begin{aligned} \min_{x} \quad & c^T x \\ \text{s.t.} \quad & x_1 A_1 + x_2 A_2 + \ldots x_n A_n \preceq B \end{aligned}$$

where $B, A_1, A_2, \ldots, A_n \in \mathbf{S}^m$ and $c \in \mathbf{R}^n$.

- Let $A(x) = x_1 A_1 + x_2 A_2 + \ldots + A_n x_n$, where $A_i \in \mathbf{R}^{q \times p}$.
- Let $\|A(x)\|_2$ be the *operator* (aka *nuclear*) norm of $A(x)$, as defined by its maximal singular value.
- Consider the following *problem of matrix norm minimization*

$$\min_x \|A(x)\|_2$$

  which is a convex problem, since $\|A(x)\|_2$ is convex in $x \in \mathbf{R}^n$.
- Using the fact that $\|A\|_2 \le t$ if and only if $A^T A \preceq t^2 I$, the above problem can be equivalently expressed as

$$\begin{aligned} \min_{x,t} \quad & t \\ \text{s.t.} \quad & A(x)^T A(x) \preceq t\, I \end{aligned}$$

- Since $A(x)^T A(x) - t\, I$ is convex in (x, t), this is a convex minimization problem with a single $(q \times q)$ matrix inequality constraint.

- We can also use the fact that, for any $t \geq 0$,

$$A(x)^T A(x) \preceq t^2\, I \iff t^2\, I - A(x)^T A(x) \succeq 0 \iff \begin{bmatrix} t\, I & A \\ A^T & t\, I \end{bmatrix} \succeq 0$$

- This results in the SDP

$$\begin{aligned} &\min_{x,t} \ \ t \\ &\text{s.t.} \ \begin{bmatrix} t\, I & A(x) \\ A(x)^T & t\, I \end{bmatrix} \succeq 0 \end{aligned}$$

- Convex optimization is a mathematically rigorous and well-studied field, with numerous numerical packages available for academic and research use.
- "With only a bit of exaggeration, we can say that, if you formulate a *practical problem* as a *convex optimization problem*, then you have solved the original problem." (S. Boyd)
- That is why it is important to recognize and formulate convex optimization problems in their *standard form*.
- Surprisingly many problems can be solved via convex optimization.