Lecture 8. Optimization problems. Linear program.

Remark. Any norm $\|\cdot\|$ is a convex function, but not strict.

Note that $\|x\| + \|3x\| = 2\|2x\|$ by homogenity, but strict requires $>$.

Verify triangle inequality for lp-norms: need generalized Cauchy-Schwarz.

Hölder's inequality. Let $p, q \in (1, \infty)$ be conjugate exponents, namely $1/p + 1/q = 1$. then $\forall u, v \in \mathbb{R}^n$ $\langle u, v\rangle \leq \|u\|_p \|v\|_q$.

Proof. Wlog assume $u_i \geq 0$, $v_i \geq 0$ and $\langle u, v\rangle > 0$. Let $\tilde{u} = u/\|u\|_p$, $\tilde{v} = v/\|v\|_q$. Our goal is to show $\Sigma |\tilde{u}_i \tilde{v}_i| = \Sigma \tilde{u}_i \tilde{v}_i \leq 1$. We claim $\forall x, y \geq 0$. $x^{1/p} y^{1/q} \leq x/p + y/q$. If $xy = 0$, trivial. Otherwise by Jensen's inequality $\log(x/p + y/q) \geq \frac{1}{p}\log x + \frac{1}{q}\log y$. Apply to $\tilde{u}_i$, $\tilde{v}_i \Rightarrow u_i v_i \leq \frac{1}{p} u_i^p + \frac{1}{q} v_i^q \Rightarrow \Sigma u_i v_i \leq \frac{1}{p}\Sigma u_i^p + \frac{1}{q}\Sigma v_i^q = 1$. $\square$

Minkowski's inequality: Let $1 \leq p \leq \infty$. $\|u+v\|_p \leq \|u\|_p + \|v\|_p$.

Proof. If $p = 1$ or ∞ or $\|u+v\|_p = 0$, trivial. Now assume $p \in (1, \infty)$.

$\|u+v\|_p^p = \Sigma |u_i + v_i|^p \leq \Sigma |u_i| \cdot |u_i + v_i|^{p-1} + \Sigma |v_i| \cdot |u_i + v_i|^{p-1}$. By Hölder's inequality, $\Sigma |u_i| \cdot |u_i + v_i|^{p-1} \leq \|u\|_p \left(\Sigma (|u_i + v_i|^{p-1})^q\right)^{1/q}$.

$\frac{1}{q} = 1 - \frac{1}{p} = \frac{p-1}{p} \Rightarrow q = \frac{p}{p-1} \Rightarrow \Sigma |u_i| \cdot |u_i + v_i|^{p-1} \leq \|u\|_p \|u+v\|_p^{p-1}$. $\square$

Convex Optimization problems.

Standard forms: $\min f(x)$ — objective function

Problem (P) subject to $g_i(x) = 0,\ i = 1, 2, \cdots, m$ — constraint functions.
$h_j(x) \le 0,\ j = 1, 2, \cdots, l$.

Domain of (P): $D \triangleq \operatorname{dom} f \cap (\bigcap_{i=1}^{m} \operatorname{dom} g_i) \cap (\bigcap_{j=1}^{l} \operatorname{dom} h_j)$.

Feasible set of (P): $\Omega \triangleq \{x \in D : \forall i.\ g_i(x) = 0,\ \forall j.\ h_j(x) \le 0\}$. feasible if $\Omega \ne \phi$

Optimal value of (P): $f^* \triangleq \inf_{x \in \Omega} f(x)$. solution: $x^* \triangleq \operatorname{argmin}_{x \in \Omega} f(x)$.

Remark. For convienence. allow f^* to take the extended value $\pm\infty$.

- $f^* = \infty$ if (P) is infeasible. i.e. $\Omega = \phi$. $\left(\begin{matrix} \sup \phi = -\infty \\ \inf \phi = \infty \end{matrix}\right)$.
- $f^* = -\infty$ if $f(x)$ is unbounded below over Ω.
- x^* is an optimal solution iff $x^* \in \Omega$ and $f(x^*) = f^*$.
- x^* is a locally optimal point if $\exists \delta > 0.\ \forall \|x - x^*\| < \delta.\ f(x) \ge f(x^*)$.

Convex Optimization: f. h_j are convex. g_i are affine (g. $-g$ convex).

Domain $D = \operatorname{dom} f \cap (\cap \operatorname{dom} h_j)$. $\Omega = \{x \in D : g_i(x) = 0,\ h_j(x) \le 0\}$.

α-sublevel set of f: $\{x \in \operatorname{dom} f : f(x) \le \alpha\}$ convex if f convex.

(α-level set: $\{x : f(x) = \alpha\}$. α-superlevel set: $\{x : f(x) \ge \alpha\}$).

$\Rightarrow \Omega$ is convex. COP: minimizing convex functions over convex sets.

Example. $\min f(x) = x_1^2 + x_2^2$ s.t. $\begin{cases} g(x) = (x_1+x_2)^2 = 0 \\ h(x) = x_1/(x_2^2+1) \le 0 \end{cases}$ not convex.

However, $f(x)$ is convex. feasible set $\Omega = \left\{ x : \begin{matrix} x_1 + x_2 = 0 \\ x_1 \le 0 \end{matrix} \right\}$ is convex.

So it is equivalent but not identical to $\min f(x)$ s.t. $\begin{cases} x_1 + x_2 = 0 \\ x_1 \le 0 \end{cases}$

Properties of convex optimization problems.

- Any local minimum is a global minimum. Trivial if unconstrained

  If $\Omega \ne \mathbb{R}^n$. $\forall x^* \in \Omega$ is a local minimum. $\forall x \ne x^* \in \Omega$. $\theta \in [0,1]$

  $x^* + \theta(x - x^*) \in \Omega \Rightarrow x - x^*$ is a feasible direction. By first-order

  condition of optimality. $\nabla f(x^*)^T(x - x^*) \ge 0 \Rightarrow f(x) \ge f(x^*)$.

- The set of optimal solutions $\Omega_{opt} = \{x^* : \forall x \ne x^*, f(x^*) \le f(x)\}$

  is convex. since $\Omega_{opt} = f^*$- sublevel set $\cap\, \Omega$ if $f^* > -\infty$. o.w. ϕ.

- In particular, if f is strictly convex. at most 1 optimal solution.

Canonical forms of convex optimization problems.

Linear program. $\min\limits_x c^T x$. s.t. $A_1 x = b_1$ $A_2 x \le b_2$.

Quadratic program. $\min\limits_x \frac{1}{2} x^T Q x + c^T x$. s.t. $A_1 x = b_1$, $A_2 x \le b_2$.

Quadratically constrained quadratic program. s.t. $\begin{matrix} \frac{1}{2} x^T Q_i x + w_i^T x + d_i \le 0 \\ Ax = b. \end{matrix}$

QP is convex if $Q \succeq 0$. QCQP is convex if $Q \succeq 0$. $\forall i$. $Q_i \succeq 0$.

Example: Linear least squares regression. Given $y \in \mathbb{R}^n$. $X \in \mathbb{R}^{n\times p}$.

goal: find $w \in \mathbb{R}^p$ to $\min\limits_w \|y - Xw\|_2^2$ $\quad (w^* = (X^TX)^{-1}X^Ty)$.

convex. $\nabla\|y-Xw\|^2 = 2X^TXw - 2y^TX$. $\nabla = 0 \Rightarrow w = (X^TX)^{-1}X^Ty$.

The simplest convex optimization problems: Linear programming.

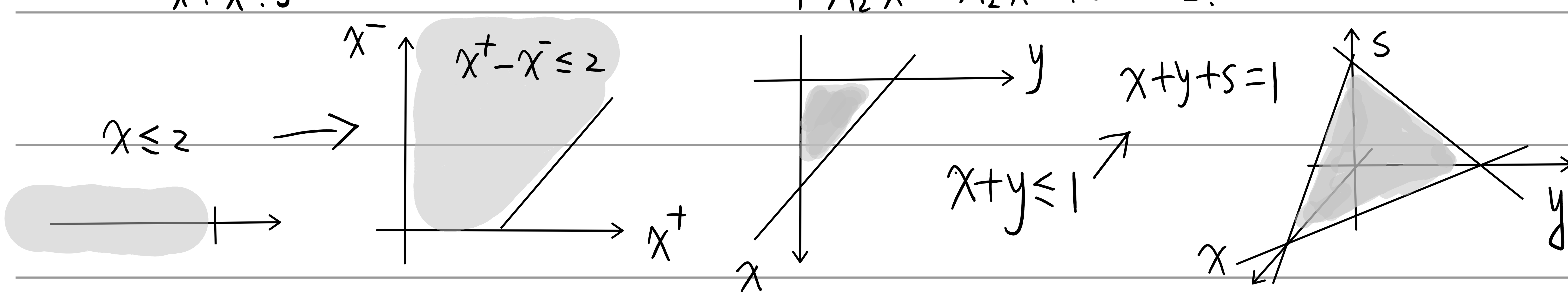
Linear program. $\min\limits_x c^Tx$ s.t. $A_1x = b_1$, $A_2x \le b_2$.

Standard form. $\min\limits_x / \max\ c^Tx$. s.t. $Ax = b$. $x \ge 0$.

- add slack variables $s_1 \dots s_m$. $\min\limits_{x,s} c^Tx$. s.t. $\begin{cases} A_1x = b_1 \\ A_2x + s = b_2. \ s \ge 0. \end{cases}$

- split variables into positive and negative parts. $x_i = x_i^+ - x_i^-$. $\forall x_i$.

$\min_{x^+, x^-, s} c^Tx^+ - c^Tx^-$. s.t. $\begin{cases} A_1x^+ - A_1x^- = b_1 \\ A_2x^+ - A_2x^- + s = b_2. \end{cases}$ $x^+, x^-, s \ge 0$.



Example: $\min -2x_1 - 3x_2$. s.t. $x_1 \le 100$. $x_2 \le 200$. $x_1 + x_2 \le 160$.

Step 1. $\min -2x_1 - 3x_2$.

s.t. $\begin{cases} x_1 + s_1 = 100 \\ x_2 + s_2 = 200 \\ x_1 + x_2 + s_3 = 160 \end{cases}$

$s_1, s_2, s_3 \ge 0$.

Step 2. $\min -2(x_1^+ - x_1^-) - 3(x_2^+ - x_2^-)$

s.t. $\begin{cases} x_1^+ - x_1^- + s_1 = 100. \\ x_2^+ - x_2^- + s_2 = 200 \\ x_1^+ - x_1^- + x_2^+ - x_2^- + s_3 = 160. \end{cases}$

$x_1^+, x_1^-, x_2^+, x_2^-, s_1, s_2, s_3 \ge 0$.

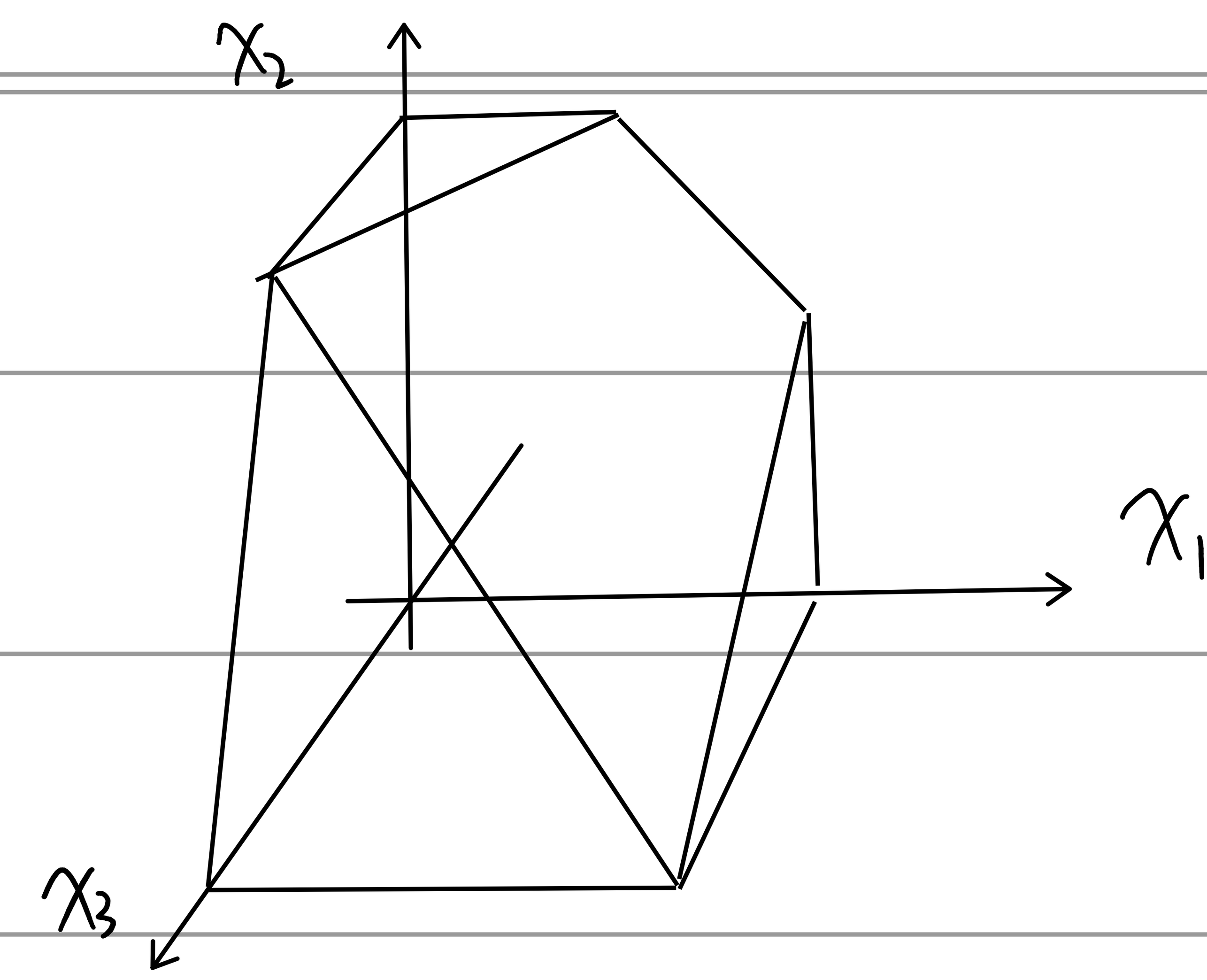
Example. $\min\ -x_1+6x_2-13x_3$

s.t. $x_1+x_2+x_3 \le 400$

$x_2+3x_3 \le 600$ $\qquad Ax \le b$

$x_1 \le 200,\ x_2 \le 300$ $\qquad x \ge 0$

$x_1\ x_2\ x_3 \ge 0$



All possibilities of LP: infeasible / unbounded / $\exists$ optimal solutions.

$\exists$ optimal: unique / infinite many (if x, y optimal, $\forall \theta x+\bar{\theta} y$ optimal).

Conjecture: if $\exists$ optimal solutions, then $\exists$ a vertex is optimal.

Definition (vertex). $x \in \mathbb{R}^n$ is a vertex of a polyhedron P, if $x \in P$ and $\exists$ n linearly independent contraints that are tight at x.

Remark. n constraints tight $\Longleftrightarrow$ n slack variable $=0$ for standard form.

Definition (Basic solution) Consider $Ax=b$, where $A \in \mathbb{R}^{m\times(m+n)}$. $\exists$ m linearly independent columns. $A=(B\in\mathbb{R}^{m\times m}, D\in\mathbb{R}^{m\times n})$, then $x=\begin{pmatrix} x_B \\ 0 \end{pmatrix}$ where $x_B\in\mathbb{R}^m$, $0\in\mathbb{R}^n$ is a basic solution. (basic feasible if $x\ge 0$).

Fundamental theorem of linear programming.

Consider the LP: max / min $c^T x$, s.t. $Ax\le b$, $x \ge 0$. Suppose the LP has ≥ 1 optimal solutions. Then $\exists$ an optimal solution at a vertex.

Algorithmic application: find optimal solutions by enumerating $\binom{m}{n}$ vertices.

However, consider an n-dimensional $[0,1]^n$ cube. only $2n$ inequalities. but 2^n vertices.