

Ch2 Convex sets

- Affine line segments**: $y = \theta x_1 + (1-\theta)x_2$, $x_1, x_2 \in E^k$, $\theta \in [0, 1]$. $\text{relint } C = \text{affine dimension} \leq 1$.
- affine sets**: $C \rightarrow \text{affine}$, $x_1, \dots, x_k \in C$, $\theta_1 + \dots + \theta_k = 1 \Rightarrow \theta_1 x_1 + \dots + \theta_k x_k \in C$, $\theta_i \in E^k$.
- affine hull**: $\text{affine hull } C = \{x \mid x = \theta_1 x_1 + \dots + \theta_k x_k, \theta_i \in E^k\}$, $\sum \theta_i = 1$.
- CE R^n** : $C = \{x \mid Ax = b\}$, $A \in E^{m,n}$, $b \in E^m$.
- affine hull of CES**: $\text{affine hull } C = \{x \mid Ax = b\}$, $A \in E^{m,n}$, $b \in E^m$.
- convex sets**: $C \rightarrow \text{convex}$, $x_1, \dots, x_k \in C$, $\theta_1 + \dots + \theta_k = 1 \Rightarrow \theta_1 x_1 + \dots + \theta_k x_k \in C$, $\theta_i \geq 0$.
- convex hull**: $\text{convex hull } C = \{x \mid \theta_1 x_1 + \dots + \theta_k x_k \in C, \theta_i \geq 0, \sum \theta_i = 1\}$.
- hyperplane & halfspaces**: $\{x \mid a^T x \leq b\}$ is affine, $\{x \mid a^T x \leq b\} \cap C$ is convex.
- Euclidean balls & ellipsoids**: $B(x_c, r) = \{x \mid \|x - x_c\|_2 \leq r\} = \{x \mid (x - x_c)^T (x - x_c) \leq r^2\} = \{x \mid x_c + ru \mid \|u\|_2 \leq 1\}$. $E = \{x \mid (x - x_c)^T P^{-1} (x - x_c) \leq 1\} = \{x \mid x_c + Au \mid \|u\|_2 \leq 1\}$. Length of $E \sqrt{\lambda} \cdot A = P^T P$.
- norm ball, norm cones**: $\text{norm ball } : C = \{x \mid \|x - x_c\|_1 \leq r\}$, $\text{norm cone } : C = \{(x, t) \mid \|x\|_1 \leq t\} \subseteq R^{n+1}$, $\text{second-order cone} : C = \{(x, z) \in E^{n+1} \mid \|x\|_1 \leq z\} = \left\{ \begin{bmatrix} x \\ z \end{bmatrix} \mid \begin{bmatrix} z & 0 \\ 0 & I_n \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} \leq 0, z \geq 0 \right\}, R^3, \{x(x, z) \mid (x^2 + z^2)^{1/2} \leq t\}$.

Operation preserve convexity

- Intersection**: $S_1, S_2 \text{ convex} \Rightarrow S_1 \cap S_2 \text{ convex}$.
- polyhedron**: intersection of halfspaces & halfplanes.
- Affine functions**: $f(x) = A^T x + b$, $f(S) = \{f(x) \mid x \in S\}$.
- geometric interpretation**: $P = \{x \mid a_j^T x \leq b_j, j=1 \dots m\}$, $a_j^T x = \sum a_{j,i} x_i, j=1 \dots p$, intersection of "FZENITTE" number of halfspaces & halfplanes.
- complex**: $C = \text{conv } \{v_0, \dots, v_k\}$, $v_0, \dots, v_k \rightarrow \text{affinely independent}$.
- Simplex**: $C = \text{conv } \{v_0, \dots, v_k\}$, $v_0, \dots, v_k \rightarrow \text{affinely independent}$.
- Simplex**: $C = \text{conv } \{v_0, \dots, v_k\}$, $v_0, \dots, v_k \rightarrow \text{affinely independent}$.
- affine hull**: $\text{affine hull } C = \{x \mid a_1 x_1 + \dots + a_k x_k \in C\}$, $a_1, \dots, a_k \in E^k$.
- CE R^n** : $C = \{x \mid Ax = b\}$, $A \in E^{m,n}$, $b \in E^m$.
- affine hull of CES**: $\text{affine hull } C = \{x \mid Ax = b\}$, $A \in E^{m,n}$, $b \in E^m$.
- convex sets**: $C \rightarrow \text{convex}$, $x_1, \dots, x_k \in C$, $\theta_1 + \dots + \theta_k = 1 \Rightarrow \theta_1 x_1 + \dots + \theta_k x_k \in C$, $\theta_i \geq 0$.
- convex hull**: $\text{convex hull } C = \{x \mid \theta_1 x_1 + \dots + \theta_k x_k \in C, \theta_i \geq 0, \sum \theta_i = 1\}$.
- hyperplane & halfspaces**: $\{x \mid a^T x \leq b\}$ is affine, $\{x \mid a^T x \leq b\} \cap C$ is convex.
- Euclidean balls & ellipsoids**: $B(x_c, r) = \{x \mid \|x - x_c\|_2 \leq r\} = \{x \mid (x - x_c)^T (x - x_c) \leq r^2\} = \{x \mid x_c + ru \mid \|u\|_2 \leq 1\}$. $E = \{x \mid (x - x_c)^T P^{-1} (x - x_c) \leq 1\} = \{x \mid x_c + Au \mid \|u\|_2 \leq 1\}$. Length of $E \sqrt{\lambda} \cdot A = P^T P$.

Operations preserve generalized inequalities

- Acute KEPⁿ, proper cone**: $K \subseteq E^k$ convex, closed, solid, nonempty interior, pointed (notline $x \in K$ $\exists x' \in K$ $x \neq x'$), proper cone induce generalized inequalities.
- Properties**: $x \leq_K y \Leftrightarrow y - x \in K$, $x \leq_K y \Leftrightarrow y - x \in \text{int } K$.
- minimum & minimal elements**: $x \in S$ every $y \in S$ $x \leq y$ no point is $x \leq y$ "more" than it.
- minimal**: $x \in S$ no point is $y \in S$ "less" than it $y \leq x$ only if $y = x$.
- minimum X**: $x \in S$ x minimum iff $S \subseteq x + K$.
- minimal X**: $x \in S$ x minimal iff $(x - K) \cap S = \{x\}$.
- Separating hyperplane theorem**: $C \cap D = \emptyset \Rightarrow \exists a^T x \leq b, a^T y \geq b$.
- LMIs (Linear matrix inequalities)**: $A(x) = x_1 A_1 + \dots + x_n A_n + B$, $B, A_i \in E^{m,m}$, $\{x \mid A(x) \leq B\} = \{x \mid f(x) = B - A(x) \leq 0\}$, $f: R^n \rightarrow S^m$, is convex.
- hyperbolic cone**: $f(x) = P^{-1} x \leq 0$, $P \in S^m$, $f(x) = P^{-1} x \leq 0$, $P \in S^m$, $f(u) = P^{-1} u \leq 0$, $\|u\|_2 \leq 1$.
- Dual generalized inequalities**: $K^* = \text{dual generalized inequality of } K$, $-x \leq_K y \iff -x^T \leq y^T \forall A \in E^{n,k} \forall x \in K, \forall y \in K^*$, $-x \leq_K y \iff -x^T \leq y^T \forall A \in E^{n,k} \forall x \in K, \forall y \in K^*$, $-x \leq_K y \iff -x^T \leq y^T \forall A \in E^{n,k} \forall x \in K, \forall y \in K^*$.
- Dual cones**: K is a cone, K^* is a dual cone, $K^* = \{y \mid x^T y \geq 0 \text{ for all } x \in K\}$.
- Linear-fractional & perspective functions**: $P: R^{n+1} \rightarrow R^n$, $\text{dom } P = R^{n+1} \setminus R_{+}$, $P(\beta, x) = \frac{x}{\beta}$.
- Linear-fractional function**: $P(C) = \{P(x) \mid x \in C\}$, convex, $P^{-1}(C) = \{x \mid P(x) \in C, x \in R^{n+1}\}$, convex.
- Linear-fractional function**: $P(x) = \begin{bmatrix} A & b \\ 0 & I_m \end{bmatrix} x + \begin{bmatrix} c \\ 0 \end{bmatrix}$, $f: R^n \rightarrow R^m$, $f = P \circ g$, $f(x) = \frac{Ax+b}{c^T x}$, $\text{dom } f = \{x \mid c^T x > 0\}$.
- Linear-fractional function**: $P(x) = \begin{bmatrix} A & b \\ 0 & I_m \end{bmatrix} x + \begin{bmatrix} c \\ 0 \end{bmatrix}$, $f: R^n \rightarrow R^m$, $f = P \circ g$, $f(x) = \frac{Ax+b}{c^T x}$, $\text{dom } f = \{x \mid c^T x > 0\}$.
- Orthogonal complement**: $V^* = \{y \mid y^T v = 0, \forall v \in V\}$.
- nonnegative without 0**: $y^T x \geq 0 \forall y \in V \setminus \{0\} \forall x \in X$.
- Self-dual**: $y^T x \geq 0 \forall y \in V \setminus \{0\} \forall x \in X$.
- Positive semidefinite cone**: $K = S^+ = \{x \in S^m \mid x^T x \geq 0, \forall x \in S^m\} \Rightarrow \text{Show } K^* = K \Rightarrow Y \in K \Rightarrow Y^T Y \geq 0$, $Y^T Y \geq 0 \Rightarrow Y \in S^m$, $Y \in S^m \Rightarrow Y^T Y \geq 0$, $Y^T Y \geq 0 \Rightarrow Y \in K$.
- Proof: self-dual**: $\text{if } Y \in K \Rightarrow Y^T Y \geq 0 \Rightarrow Y^T Y \geq 0 \Rightarrow Y \in S^m$, $\text{if } Y \in S^m \Rightarrow Y^T Y \geq 0 \Rightarrow Y \in K$, $\Rightarrow Y^T Y \geq 0 \Rightarrow Y \in K$.
- affine**: $x_1, \dots, x_k \in E^k$, $x_1, \dots, x_k \rightarrow \text{affinely independent}$.
- convex**: $x_1, \dots, x_k \in E^k$, $x_1, \dots, x_k \rightarrow \text{convex}$.
- closed**: $x_1, \dots, x_k \in E^k$, $x_1, \dots, x_k \rightarrow \text{closed}$.
- proper**: $x_1, \dots, x_k \in E^k$, $x_1, \dots, x_k \rightarrow \text{proper}$.
- closed & convex**: $x_1, \dots, x_k \in E^k$, $x_1, \dots, x_k \rightarrow \text{closed & convex}$.
- convex & closed**: $x_1, \dots, x_k \in E^k$, $x_1, \dots, x_k \rightarrow \text{convex & closed}$.
- proper & closed**: $x_1, \dots, x_k \in E^k$, $x_1, \dots, x_k \rightarrow \text{proper & closed}$.
- proper & convex**: $x_1, \dots, x_k \in E^k$, $x_1, \dots, x_k \rightarrow \text{proper & convex}$.
- closed & proper**: $x_1, \dots, x_k \in E^k$, $x_1, \dots, x_k \rightarrow \text{closed & proper}$.
- convex & proper**: $x_1, \dots, x_k \in E^k$, $x_1, \dots, x_k \rightarrow \text{convex & proper}$.
- closed & convex & proper**: $x_1, \dots, x_k \in E^k$, $x_1, \dots, x_k \rightarrow \text{closed & convex & proper}$.

Ch3 convex functions

(Definition)

$$f: \mathbb{R}^n \rightarrow \mathbb{R}$$

if dom f is convex set
 $x, y \in \text{dom } f \in [0, 1]$
 $\Rightarrow f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$

\Rightarrow f convex, f concave

\Rightarrow f is convex, i.e. f.

$g(t) = f(x+tV)$ is convex
 $\Rightarrow f(x+tV) \in \text{dom } f$ (line)

(1st-order condition)

f differentiable

f convex i.f.f.

dom f convex

$\Rightarrow f'(x) \geq 0$ (convex)

(concave)

$f''(x) \geq 0$ (positive)

(semi-definite)

(concave)

(Examples)

Exponential e^x on \mathbb{R} convex

Power x^α on $\mathbb{R} \setminus \{0\}$ convex

x^α -- quasi convex

Power of absolute value

$|x|^p$, $p \geq 1$

Logarithm $\log x$ convex

Negative entropy $x \log x$ convex

Norms convex

Max function convex

$f(x) = \max\{x_1, \dots, x_n\}$

Quadratic over-linear function

$f(x) = x_1^2/x_2$, $x \in \mathbb{R}^n, x_2 > 0$ convex

Log-sum-exp

$f(x) = \log(e^{x_1} + \dots + e^{x_n})$ convex

Geometric mean

$f(x) = (\prod_{i=1}^n x_i)^{1/n}$ convex

Log-determinant

$f(x) = \log \det(X)$ on $\text{dom } f = S_{++}^n$

convex

Method in sum:

1. Check basic inequality

2. 2nd-order: Hessian Matrix

3. restrict to an arbitrary line

& Verify convexity on R

e.g. $g(t) = \log \det(I/t + V)$

4. operations

(Sublevel sets)

α -sublevel set: $f: \mathbb{R}^n \rightarrow \mathbb{R}$

$C_\alpha = \{x \in \text{dom } f | f(x) \leq \alpha\}$ (if convex function is convex)

α -superlevel set:

$C_\alpha = \{x \in \text{dom } f | f(x) \geq \alpha\}$ (if convex function is convex)

(Epigraph)

graph of $f: \mathbb{R}^n \rightarrow \mathbb{R}$

$\equiv \{(x, t) | x \in \text{dom } f\} \subseteq \mathbb{R}^{n+1}$

Epigraph of $f: \mathbb{R}^n \rightarrow \mathbb{R}$ convex

$\equiv \{(x, t) | x \in \text{dom } f, t \geq f(x)\} \subseteq \mathbb{R}^{n+1}$

\Rightarrow a function is convex i.f.f.

epigraph is convex set.

Hypograph of $f: \mathbb{R}^n \rightarrow \mathbb{R}$ convex

$\equiv \{(x, t) | x \in \text{dom } f, t \leq f(x)\} \subseteq \mathbb{R}^{n+1}$

(Jensen's inequality & extensions)

$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$

$\Rightarrow f(\theta x_1 + \dots + \theta x_k) \leq \theta f(x_1) + \dots + \theta f(x_k), \forall \theta \in [0, 1]$

$\Rightarrow f(\int_S p(x)dx) \leq \int_S f(x)p(x)dx$

$\Rightarrow f(E_x) \leq E f(x)$

$\Rightarrow \Rightarrow$

Convex inequality:

$\text{prob}(x=x_1) = \theta, \text{prob}(x=x_2) = (1-\theta)$

$\therefore f(\theta x_1 + (1-\theta)x_2) \leq \theta f(x_1) + (1-\theta)f(x_2)$

(Cauchy-Schwarz inequality)

$(a^T a)(b^T b) \geq (a^T b)^2$

Operations that preserve convexity

Nonnegative weighted sum

$\Rightarrow f = w_1 f_1 + \dots + w_m f_m$

is convex - given

f_1, \dots, f_m are convex

Composition w/ affine mapping

$\Rightarrow f: \mathbb{R}^n \rightarrow \mathbb{R}$

$A \in \mathbb{R}^{n \times m}, b \in \mathbb{R}^m$

$g(x) = f(Ax+b)$, $\text{dom } g$

$\subseteq \{x | Ax+b \in \text{dom } f\}$

\Rightarrow if f convex, g is convex

$\Rightarrow f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$

$\Rightarrow f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$

$\Rightarrow f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$

$\Rightarrow f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$

$\Rightarrow f(x) = \max\{f_1(x), \dots, f_m(x)\}$

e.g. $f(x) = \sum_{i=1}^n x_i$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n\}$

$\Rightarrow f(x) = \sum_{i=1}^n x_i$

$= \max\{x_1, \dots, x_n$

Ch4 Convex problem

Basic terminology

minimize $f_0(x)$
subject to $f_i(x) \leq 0, i=1, \dots, m$
 $h_i(x) = 0, i=1, \dots, p$

(optimal value)

$p^* = \inf\{f_0(x) | f_i(x) \leq 0, i=1, \dots, m$
 $\quad \quad \quad h_i(x) = 0, i=1, \dots, p\}$

$p^* = \infty$ if problem infeasible
 $p^* = -\infty$ if problem unbounded below

D.T.W. $D = \bigcap_{i=1}^m \text{dom } f_i \cap \bigcap_{i=1}^p \text{dom } h_i$

(optimal & locally optimal points)

$\rightarrow X_{\text{opt}} = \{x | f_0(x) \leq 0, h_i(x) = 0, f_i(x) = p^*, i=1, \dots, m, i=h, \dots, p\}$

$\rightarrow X @ f_0(x) \leq p^* + \epsilon, \epsilon \text{-suboptimal}$

$\rightarrow f_0(x) = \inf\{f_0(y) | f_i(y) \leq 0, h_i(y) = 0$
 $\quad \quad \quad \text{if } \exists R > 0 \quad \|y-x\|_2 \leq R\}$ (locally problem)

\rightarrow optimal value attained
achieved
infeasible \Rightarrow solvable
unbounded \Rightarrow ∞

find x subject to $f_i(x) \leq 0, h_i(x) = 0$
 \rightarrow feasibility problem

implicit constraints $x \in D = \bigcap_{i=1}^m \text{dom } f_i \cap \bigcap_{i=1}^p \text{dom } h_i$
explicit constraints $h_i(x) = 0, f_i(x) \leq 0$

Convex Optimization

minimize $f_0(x)$
subject to $f_i(x) \leq 0, i=1, \dots, m$ convex
 $a_i^T x = b_i, i=1, \dots, p$ affine

feasible set $D = \bigcap_{i=0}^m \text{dom } f_i$

\rightarrow minimize a convex objective function
over
a convex set

Local & global optima

locally optima = global optima

$\rightarrow x$ is locally optimal $\Leftrightarrow x$ feasible

$f_0(x) = \inf\{f_0(y) | y \text{ feasible}, \|y-x\|_2 \leq R\}$
 $R > 0$

\rightarrow proof:

if x not globally optimal,

$\exists y \quad f_0(y) < f_0(x), \|y-x\|_2 > R$

also

$$g = (1-\theta)x + \theta y \quad \theta = \frac{R}{\|y-x\|_2} \rightarrow \|g-x\|_2 = \frac{R}{2} < R$$

$$\therefore f_0(g) \leq (1-\theta)f_0(x) + \theta f_0(y) < \theta f_0(x)$$

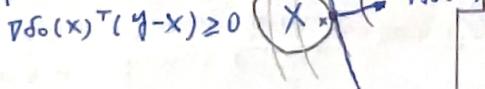
↑ contradicts

$$f_0(x) = \inf\{f_0(y) | y \text{ feasible}, \|y-x\|_2 \leq R\}$$

(optimality criterion)

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

x is optimal $\Leftrightarrow \nabla f_0(x) = 0$



more on vector optimization

Scalarization & \mathbb{R}^n
for any $\lambda \neq 0$, if \tilde{x} is an optimal point
for the scalar optimization problem below

minimize $\lambda^T f_0(x)$
subject to $f_i(x) \leq 0, i=1, \dots, m$
 $h_i(x) = 0, i=1, \dots, p$

\rightarrow then \tilde{x} is pareto optimal for the
vector optimization problem

\rightarrow for every pareto optimal point x^* ,

$\exists \lambda \geq 0, \lambda \neq 0$, such that

\tilde{x} is an optimal point of
scalarized problem

(equivalent convex problems)

(Eliminating equality constraints)

minimize $f_0(x)$
subject to $f_i(x) \leq 0, i=1, \dots, m$
 $h_i(x) = 0, i=1, \dots, p$

\downarrow
minimize $f_0(Fg+x_0)$
subject to $f_i(Fg+x_0) \leq 0, i=1, \dots, m$

\downarrow
introducing equality constraints
minimize $f_0(Ax+b_0)$
subject to $f_i(Ax+b_i) \leq 0, i=1, \dots, m$

\downarrow
minimize $f_0(y)$
subject to $f_i(y) \leq 0, i=1, \dots, m$
 $y = Ax+b_i, i=0, 1, \dots, m$

(introducing slack variables for linear inequalities)

minimize $f_0(x)$
subject to $a_i^T x \leq b_i, i=1, \dots, m$

\downarrow
minimize $f_0(x)$
subject to $A_i^T x + s_i = b_i, i=1, \dots, m$
 $s_i \geq 0, i=1, \dots, m$

(epigraph problem form)

minimize t
 $x+t$
subject to $f_0(x)-t \leq 0$
 $f_i(x) \leq 0, i=1, \dots, m$
 $a_i^T x + b_i \leq t, i=1, \dots, p$

(minimizing over some variables)

minimize $f_0(x_1, x_2)$
subject to $f_i(x_1) \leq 0, i=1, \dots, n$

\downarrow
minimize $\tilde{f}_0(x_1)$
subject to $f_i(x_1) \leq 0, i=1, \dots, m$
where $\tilde{f}_0(x_1) = \inf f_0(x_1, x_2)$

(Quasiconvex function)

minimize $f_0(x)$
subject to $f_i(x) \leq 0, i=1, \dots, m$
 $AX=b$

w/ $f_0: \mathbb{R}^n \rightarrow \mathbb{R}$ quasiconvex
 f_1, \dots, f_m convex

$f_0(x) \leq t \Leftrightarrow \partial f_0(x) \leq 0$
 t -sublevel
 $\subseteq \mathbb{R}^n$

\downarrow
formulate as
finding x
subject to $\partial f_0(x) \leq 0$
 $\rightarrow f_i(x) \leq 0, i=1, \dots, m$
 $AX=b$

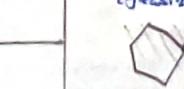
\rightarrow suppose f_0 is differentiable
Let \bar{X} be the feasible set.
if $x \in \bar{X} \wedge \nabla f_0(x)^T(y-x) > 0$

$\forall y \in \bar{X} \setminus \{x\} :$
 x is optimal

(Linear Optimization Problem)

minimize $C^T x + d$ affine
subject to $Gx \leq h$ affine
 $\rightarrow AX=b$ affine

(feasible set polyhedron)



Linear-fractional program

minimize $f_0(x)$
subject to $Gx \leq h$
 $AX=b$

\downarrow
 $f_0(x) = \frac{C^T x + d}{B^T x + b}$ dom $f_0 = \{x | B^T x + b > 0\}$

is equivalent

\downarrow
minimize $C^T x + d$
subject to $Gy \leq h$
 $Ay = b$
 $B^T y + b = 1$
 $\exists z \geq 0$

(Quadratic Program)

minimize $\frac{1}{2} x^T P x + q^T x + r$
subject to $Gx \leq h$
 $AX=b$

PES⁺
 \rightarrow $f_0(x^*) = \min f_0(x)$

\downarrow
proper cone

(Generalized inequality constraints)

minimize $f_0(x) : \mathbb{R}^n \rightarrow \mathbb{R}$
subject to $f_i(x) \leq k_i, i=1, \dots, m$
 $AX=b$

\downarrow
proper cone

\rightarrow convex feasible set

\rightarrow local optima
global
differentiable,
constraint holds

\rightarrow $AX=b$

\rightarrow semidefinite programming

$K \in S^k$ LMI
minimize $C^T x$
subject to $x_i F_i + \dots + x_n F_n \leq 0$
 $\rightarrow AX=b$

\rightarrow LP \rightarrow SDP

minimize $C^T x + d$
subject to $Gx \leq h$
 \downarrow
 $AX=b$

\rightarrow SDP \rightarrow LP

minimize $\frac{1}{2} x^T P x + q^T x + r$
subject to $Gx \leq h$
 \downarrow
 $AX=b$

\rightarrow LP \rightarrow SDP

minimize $\frac{1}{2} x^T P x + q^T x + r$
subject to $Gx \leq h$
 \downarrow
 $AX=b$

\rightarrow SDP \rightarrow LP

minimize $\frac{1}{2} x^T P x + q^T x + r$
subject to $Gx \leq h$
 \downarrow
 $AX=b$

\rightarrow LP \rightarrow SDP

minimize $\frac{1}{2} x^T P x + q^T x + r$
subject to $Gx \leq h$
 \downarrow
 $AX=b$

\rightarrow SDP \rightarrow LP

minimize $\frac{1}{2} x^T P x + q^T x + r$
subject to $Gx \leq h$
 \downarrow
 $AX=b$

\rightarrow LP \rightarrow SDP

minimize $f^T x$
subject to $\|Ax-b\|_2 \leq c_i^T x + d_i$
 \downarrow
 $FX=g$

\rightarrow SDP \rightarrow LP

minimize $\frac{1}{2} x^T P x + q^T x + r$
subject to $\|Ax-b\|_2 \leq c_i^T x + d_i$
 \downarrow
 $FX=g$

\rightarrow LP \rightarrow SDP

minimize $f^T x$
subject to $\|Ax-b\|_2 \leq c_i^T x + d_i$
 \downarrow
 $FX=g$

\rightarrow SDP \rightarrow LP

minimize $\frac{1}{2} x^T P x + q^T x + r$
subject to $\|Ax-b\|_2 \leq c_i^T x + d_i$
 \downarrow
 $FX=g$

\rightarrow LP \rightarrow SDP

\rightarrow SDP \rightarrow LP

Algorithms (cont'd) Ch 9, 10, 11 Ch 9-10, 11

Justification of stopping criterion

$$\begin{aligned} f(x) - p^* \\ = f(x) - \inf_{y \in \mathbb{R}^n} \{f(y) | Ay = b\} \\ \Rightarrow f(x) - \inf_{y \in \mathbb{R}^n} \{\delta(y) | Ay = b\} \\ = f(x) - \inf_{v \in \mathbb{R}^n} \{f(x+v) | A(x+v) = b\} \\ = f(x) - \inf_{v \in \mathbb{R}^n} \{f(x+v) | Av = 0\} \\ = f(x) - \inf_{v \in \mathbb{R}^n} \{f(x) + \nabla f(x)^T v + \frac{1}{2} v^T \nabla^2 f(x) v\} \\ Av = 0 \end{aligned}$$

solve ① = solve ②

Newton's method for equality constraints

Derivation of Newton Step

$$\begin{aligned} \text{① } @x, \\ \text{addition of } v: x+v \\ \text{x}+v \text{ should be feasible} \\ A(x+v) = b \\ Av = 0 \\ \text{② } f(x+v) = f(x) + \nabla f(x)^T v + \frac{1}{2} v^T \nabla^2 f(x) v \\ \text{from ① \& ② we get} \\ \text{minimize } f(x) + \nabla f(x)^T v + \frac{1}{2} v^T \nabla^2 f(x) v \\ \text{subject to } Av = 0 \quad \text{③} \\ \text{③ } v \text{ is optimal iff. } \exists w, \text{ such that } (v, w) \text{ satisfies KKT!} \end{aligned}$$

$$\begin{aligned} \text{④ } \begin{cases} A\tilde{v} = 0 \\ \nabla f(x) + \nabla^2 f(x)\tilde{v} + A^T w = 0 \end{cases} \\ \Leftrightarrow \begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \tilde{v} \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix} \\ \Delta x_{nt} = \tilde{v} \quad \text{KKT matrix.} \end{aligned}$$

$$\Leftrightarrow \begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{nt} \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}$$

- Derivation of New Step (Again!)

$$\begin{aligned} \text{⑤ } @x \\ \text{addition of } v: x+v \\ (x+v, w) \text{ satisfies ③ (KKT)!} \\ \begin{cases} A(x+v) = b \\ \nabla f(x+v) + A^T w = 0 \end{cases} \quad \text{To do: linearize} \\ \Leftrightarrow \begin{cases} A(x+v) = b \\ \nabla f(x) + \nabla^2 f(x)v + A^T w = 0 \end{cases} \quad \text{written} \\ \Leftrightarrow \begin{cases} Av = 0 \\ \nabla f(x) + \nabla^2 f(x)v + A^T w = 0 \end{cases} \\ \Leftrightarrow \begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix} \\ \Delta x_{nt} = \tilde{v} \\ \Leftrightarrow \begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{nt} \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix} \end{aligned}$$

The Newton decrement

- pretty similar to unconstrained case

$$-\lambda(x) = (\Delta x_n^T \nabla^2 f(x) \Delta x_n)^{0.5}$$

- descent direction

$$\text{- from ④} \\ \nabla f(x) \Delta x_{nt} + A^T w = -\nabla f(x)$$

- for descent direction

$$\nabla f(x)^T \Delta x_{nt} < 0$$

$$\Rightarrow -(\nabla^2 f(x) \Delta x_{nt} + A^T w) \Delta x_{nt} < 0$$

$$= -\Delta x_n^T \nabla^2 f(x) \Delta x_{nt} + A^T w \Delta x_{nt} < 0$$

$$= -\Delta x_n^T \nabla^2 f(x) \Delta x_{nt} + 0$$

$$= -\lambda^2(x)$$

$$< 0$$

- we at central point

- we got dual feasible at $x^*(t)$

with

$\lambda^*(t), \nu^*(t)$

convergence analysis

similar to unconstrained

Interior Point Method

$$\begin{aligned} \text{minimize } f_0(x) \quad \text{①} \\ \text{subject to } f_i(x) \leq 0 \quad i=1 \dots m \\ Ax = b \quad A \in \mathbb{R}^{m \times n} \quad \text{rank}(A) = P \end{aligned}$$

- assume strictly feasible
 $\exists x \in D$ such that $f_i(x) < 0 \quad i=1 \dots m$

$$Ax = b$$

- transfer inequality to equality

Logarithmic Barrier

$$I-(u) = \begin{cases} 0 & \text{if } u \leq 0 \\ -\infty & \text{if } u > 0 \end{cases}$$

$$\text{minimize } f_0(x) + \sum_{i=1}^m I-(f_i(x)) \quad \text{such that } Ax = b \quad \text{②}$$

$$\text{define } \phi(x) = -\sum_{i=1}^m \log(-f_i(x))$$

- where:

$$\nabla \phi(x) = \sum_{i=1}^m \frac{1}{-f_i(x)} \nabla f_i(x)$$

$$\nabla^2 \phi(x) = \sum_{i=1}^m \frac{1}{-f_i(x)^2} \nabla f_i(x) \nabla f_i(x)^T + \sum_{i=1}^m \frac{1}{-f_i(x)} \nabla^2 f_i(x)$$

$$\text{minimize } t f_0(x) + \phi(x) \quad \text{③}$$

$$\text{subject to } Ax = b \quad \phi(x) = -\sum_{i=1}^m \log(-f_i(x))$$

Central Path

$$-\frac{1}{t} \phi(x) + t \uparrow, \text{ closer to } I-(u)$$

$$-\text{optimal solution of ③} \quad x^*(t) \quad \text{central point}$$

$$-\text{central path: } \{x^*(t) | t > 0\}$$

- Question:

How good is $x^*(t)$?

A. minimize $t f_0(x) + \phi(x)$

subject to $Ax = b$

• $x^*(t) \rightarrow$ strictly feasible:

$$Ax^*(t) = b \quad f_i(x^*(t)) < 0$$

• $\exists \lambda \in \mathbb{R}^P$ such that

$$+ \nabla f_0(x^*(t)) + \nabla \phi(x^*(t)) + A^T \lambda = 0$$

$$\Rightarrow + \nabla f_0(x^*(t)) + \sum_{i=1}^m \frac{1}{-f_i(x^*(t))} \nabla f_i(x^*(t)) + A^T \lambda = 0$$

$$\Rightarrow \nabla f_0(x^*(t)) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*(t)) + A^T \lambda^* = 0$$

$$\Rightarrow x^*(t) = \arg \min L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x^*(t)) + \nu^* (Ax - b)$$

$$\Rightarrow \phi(x^*(t), \nu^*(t)) = f_0(x^*(t)) + \sum_{i=1}^m \lambda_i^* f_i(x^*(t)) + \nu^* (t)^T (A x^*(t) - b)$$

$$= f_0(x^*(t)) + \sum_{i=1}^m \frac{1}{-f_i(x^*(t))} f_i(x^*(t)) + 0 = f_0(x^*(t)) - \frac{m}{t}$$

hyperbolic constraints

$$- x^T x \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- x^T x \leq 4y^2 \quad y \geq 0 \quad y \geq 0$$

$$- x^T x + y^2 + z^2 \leq y^2 + z^2 + 4y^2 \quad y \geq 0 \quad y \geq 0$$

$$- 4x^T x + y^2 + z^2 - 2y^2 \leq y^2 + z^2 + 2y^2 \quad y \geq 0 \quad y \geq 0$$

$$- 4x^T x + (y-z)^2 \leq (y+z)^2 \quad y \geq 0 \quad y \geq 0$$

$$- 4x^T x + (y-z)^2 \leq (y+z)^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \quad y \geq 0$$

$$- ||[x^T]||_2^2 \leq y^2 \quad y \geq 0 \$$

Linear Algebra Basics :

- 4 ways to write a linear system
- $\begin{cases} 2x_1 + 3x_2 - 2x_3 = 7 \\ x_1 - x_2 - 3x_3 = 5 \end{cases}$ system of equations
- $\left(\begin{array}{ccc|c} 2 & 3 & -2 & 7 \\ 1 & -1 & -3 & 5 \end{array} \right)$ augmented matrix
- $x_1 + x_2 + x_3 = b$
- $x_1 + x_2 + x_3 = b$ vector equation
- $\begin{pmatrix} 2 & 3 & -2 \\ 1 & -1 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$
- $Ax = b$ \exists soln i.f.f. b is in the span of column of A

Schurz complement

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succeq 0$$

$$\downarrow$$

$$C - B^T A^{-1} B \succeq 0$$

Matrix manipulation

$$\begin{aligned} (AB)^{-1} &= B^{-1}A^{-1} \\ (ABC)^{-1} &= C^{-1}B^{-1}A^{-1} \\ (A^T)^{-1} &= (A^{-1})^T \\ (A+B)^T &= A^T + B^T \\ (AB)^T &= B^T A^T \\ (ABC)^T &= -C^T B^T A^T \end{aligned}$$

$$\text{tr}(A) = \sum_i \text{A}_{ii}$$

$$\text{tr}(A) = \sum_i \lambda_i; \lambda_i = \text{eig}(A)$$

$$\text{tr}(A) = \text{tr}(A^T)$$

$$\text{tr}(AB) = \text{tr}(BA)$$

$$\text{tr}(A+B) = \text{tr}(A) + \text{tr}(B)$$

$$\text{tr}(ABC) = \text{tr}(BCA) = \text{tr}(CAB)$$

$$\text{tr}(AA^T) = \text{tr}(AAT)$$

$$\det(A) = \prod_i \lambda_i; \lambda_i = \text{eig}(A)$$

$$\det(cA) = c \det(A)$$

$$\det(A^T) = \det(A)$$

$$\det(AB) = \det(A)\det(B)$$

$$\det(A^{-1}) = 1/\det(A)$$

$$\det(A^n) = \det(A)^n$$

$$\det(I+uv^T) = 1+u^T v$$

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix}$$

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \dots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$$

PD matrix M

$\text{eig}(M) > 0$, symmetric

convert QP to SOCP

$$\begin{aligned} \text{minimize} \quad & x^T A x + a^T x \\ \text{subject to} \quad & Bx \leq b \\ & s.t. \end{aligned}$$

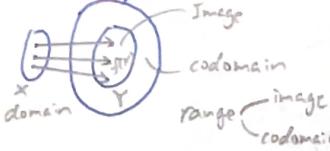
$$\begin{aligned} \Rightarrow \text{minimize} \quad & y + a^T x \\ \text{s.t.} \quad & Bx \leq b \\ & x^T A x \leq y \end{aligned} \Rightarrow \begin{aligned} & y \geq x^T A x \\ & \Rightarrow 0 \geq x^T A x - y \\ & \Rightarrow 0 \geq 4x^T A x - 4y \\ & \Rightarrow 0 \geq 4x^T A x + (1-y)^2 - (1+y)^2 \\ & \Rightarrow (1+y)^2 \geq 4x^T A x + (1-y)^2 \end{aligned}$$

Grammatic

$$\begin{bmatrix} a & b & c \\ a^2 & ab & ac \\ b^2 & bc & bc \\ c^2 & ac & bc \end{bmatrix}$$

for R^3

domain, range



Integration by change of variables.

$$\int_0^x f(t) dt \quad t=sx \quad \int_{\theta(b)}^{f(b)} f(u) du \quad u=\theta(x)$$

$$\int_a^b f(sx) ds \quad \int_a^b f(g(x)) g'(x) dx$$

$$\Delta B(x, \epsilon) = \{y \in R^n | \|y-x\|_2 \leq \epsilon\}$$

interior point: $\exists \epsilon > 0$

$$x \in \text{int } B(x, \epsilon) \subset C$$

LO, Li-Yu

20997405

llloac@connect.ust.hk

$$\textcircled{O} \setminus \textcircled{x} \rightarrow \text{int } C$$

limit point of C

if $\forall \epsilon > 0$, excluding

$$(B(x, \epsilon) \setminus \{x\}) \cap C \neq \emptyset$$

or

x is limit point of set S

if $\forall \epsilon > 0, \exists y \in S \setminus \{x\}$

$$w \delta(x, y) < \epsilon$$

closure

$$cl(C) = C \cup L(C) \quad \text{includes } bd(C)$$

cl(C), closed

cl(C), smallest closed set

contains C

$$C \subseteq S,$$

$$cl(C) \subseteq S$$

set C is closed i.f.f.

$$C = cl(C)$$

Boundary

$$bd(C) = cl(C) \setminus \text{int}(C)$$

$$\text{int}(C) \subseteq C \subseteq cl(C)$$

$$C \text{ open i.f.f. } C \cap bd(C) = \emptyset$$

$$C \text{ closed i.f.f. } bd(C) \subseteq C$$

$$\Delta f_0(x) = \frac{1}{2} x^T P x + q^T x + r_0$$

$$\nabla f_0(x) = \frac{1}{2} (P + P^T) x + q$$

$$\Delta \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{bmatrix} \quad \text{eg. of SPD}$$

$\Delta \det[\lambda I - A]$ polynomial \rightarrow find eigenvalues

$\det[\lambda B - A] \quad \dots \rightarrow$ find generalized eigenvalues

linear-fractional programming \rightarrow quasiconvex function

$$\begin{aligned} \text{minimize} \quad & f(x) = \frac{c^T x + d}{e^T x + f} \quad \text{dom } f = \{x | e^T x + f > 0\} \\ \text{subject to} \quad & Gx \leq h \\ & Ax = b \\ & x \geq 0 \end{aligned}$$

$$\text{if } \{x | Gx \leq h, Ax = b, e^T x + f > 0\} \neq \emptyset$$

$$\begin{aligned} \text{minimize} \quad & C^T y + d \bar{y} \\ \text{subject to} \quad & Gy - h \bar{y} \leq 0 \\ & Ay - b \bar{y} = 0 \\ & \bar{y}^T y + f \bar{y} = 1 \\ & \bar{y} \geq 0 \end{aligned} \quad \begin{aligned} y &= \frac{x}{e^T x + f} \\ \bar{y} &= \frac{1}{e^T x + f} \end{aligned}$$

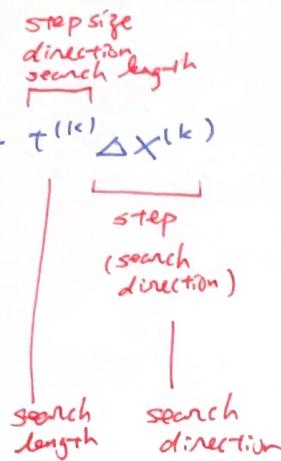
minimize $y + a^T x$

$$\text{s.t.} \quad \left\| \begin{bmatrix} 2A^T x \\ 1-y \end{bmatrix} \right\|_2 \leq 1+y$$

$$Bx \leq b$$

Descent Method

$$\Delta x^{(k+1)} = x^{(k)} + t^{(k)} \Delta x^{(k)}$$



$$\Delta f(x^{(k+1)}) < f(x^{(k)})$$

Δ General Descent method

1. (given) $x^{(0)} \in \text{dom } f$

2. (repeat)

a. get descent direction Δx

$\Delta x^{(k)}$ Search direction

b. line search $t^{(k)}$ —————— search length

1. Exact Line Search

$$t = \underset{s \geq 0}{\operatorname{argmin}} f(x + s\Delta x)$$

c. update

$$x := x + t\Delta x$$

3. (until)

$$\|\nabla f(x)\|_2 \leq \eta \quad (\text{oftenly})$$

Gradient Descent Method

1. (given) $x^{(0)} \in \text{dom } f$

2. (repeat)

a. $\Delta x^{(k)} = -\nabla f(x^{(k)})$

b. backtracking line search get $t^{(k)}$

c. $x^{(k+1)} = x^{(k)} + \Delta x^{(k)}$

3. (until)

$$\|\nabla f(x)\|_2 \leq \eta$$

2. $t := 1$

3. (while)

$$f(x + t\Delta x) > f(x) + \alpha t \nabla f(x)^T \Delta x$$

(do)

$$t := \beta t$$

Newton's Method

1. (given) $x^{(0)} \in \text{dom } f$

2. (repeat)

a. $\Delta x^{(k)} := -\nabla^2 f(x^{(k)})^{-1} \nabla f(x^{(k)})$

$$x^{(k+1)} := x^{(k)} + \Delta x^{(k)}$$

b. (until) $\frac{\|x^{(k+1)} - x^{(k)}\|_2}{2} \leq \epsilon$

c. backtracking line search get $t^{(k)}$

d. $x^{(k+1)} = x^{(k)} + t^{(k)} \Delta x^{(k)}$

Newton's method

equality constrained

1. (given) $x^{(0)} \in \text{dom } f$
 $Ax = b$

2. (repeat)

a. $\Delta x_{nt}^{(k)} := \begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix}^{-1} \begin{bmatrix} -\nabla f(x) \\ w \end{bmatrix} \quad (1:n, :)$
 $\lambda^{(k)^2} := \nabla f(x^{(k)})^T \nabla^2 f(x^{(k)})^{-1} \nabla f(x^{(k)})$

b. (until) $\lambda^{(k)^2}/2 \leq \epsilon$

c. backtracking line search get
 $t^{(k)}$

d. $x^{(k+1)} = x^{(k)} + t^{(k)} \Delta x^{(k)}$

Newton's method (infeasible start)

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{pd} \\ \Delta \nu_{pd} \end{bmatrix} = - \begin{bmatrix} \boldsymbol{\pi}_{\text{dual}} \\ \boldsymbol{\pi}_{\text{pri}} \end{bmatrix} = - \begin{bmatrix} \nabla f(x) + A^T \nu \\ Ax - b \end{bmatrix} \quad \textcircled{+}$$

1. (given) $x^{(0)} \in \text{dom } f$

ν

$G > 0$

$\alpha \in (0, 0.5)$

$\beta \in (0, 1)$

2. (repeat)

a. $\Delta x_{nt}^{(k)} := \text{from } \textcircled{+}$

$\Delta \nu_{nt}^{(k)} := \text{from } \textcircled{+}$

b. backtracking line search on $\|\boldsymbol{\pi}\|_2$

1. $t := 1$

2. while $\|\boldsymbol{\pi}(x + t \Delta x_{nt}, \nu + t \Delta \nu_{nt})\|_2 > (1-\alpha t) \|\boldsymbol{\pi}(x, \nu)\|_2$

$t := \beta t$

c. $x := x + t \Delta x_{nt}$

$\nu := \nu + t \Delta \nu_{nt}$

3. (until)

$Ax = b$

$\|\boldsymbol{\pi}(x, \nu)\|_2 \leq \epsilon$

Barrier method with logarithm

inequality constrained

original problem:

$$\begin{aligned} & \text{minimize } f_0(x) \\ & \text{subject to } f_i(x) \leq 0 \quad i=1, \dots, m \\ & Ax = b \end{aligned}$$



new problem:

$$\text{minimize } T f_0(x) + \phi(x)$$

$$\text{subject to } Ax = b$$

$$\phi(x) = -\sum_{i=1}^m \log(-f_i(x))$$

1. (given) $x^{(0)}$ (feasible)

$$T^{(0)} := T^{(0)} > 0$$

$$\mu > 1$$

$$\epsilon > 0$$

2. (repeat)

1. solve $x^*(\tau)$ of $T f_0 + \phi$ subject to $Ax = b$
with " τ "

$$2. x := x^*(\tau)$$

3. (until)

$$\frac{\|x\|}{\tau} < \epsilon$$

$$4. \tau := \mu \tau$$

$$\nabla \phi(x) = \sum_{i=1}^m -\frac{1}{-f_i(x)} \nabla f_i(x)$$

$$\nabla^2 \phi(x) = \sum_{i=1}^m \frac{1}{f_i(x)^2} \nabla f_i(x) \nabla f_i(x)^T + \sum_{i=1}^m \frac{1}{-f_i(x)} \nabla^2 f_i(x)$$

Phase I via infeasible start Newton method

inequality constrained
w/ infeasible
start.

original problem:

$$\text{minimize } f_0(x)$$

$$\text{subject to } f_i(x) \leq 0 \quad i=1, \dots, m$$

$$Ax = b$$



$$\text{minimize } f_0(x)$$

$$\text{subject to } f_i(x) \leq S, \quad i=1, \dots, m$$

$$Ax = b$$

$$S = 0$$



$$\text{minimize } t^{10} f_0(x) - \sum_{i=1}^m \log(S - f_i(x))$$

$$\text{subject to } Ax = b$$

$$S = 0$$