Basic phase I method.

Find x such that

$$f_i(x) < 0, \quad i=1,\dots,m. \quad Ax=b. \quad (2)$$

– phase I optimization $\boxed{\text{variables } x, s}$

$$\begin{aligned} &\text{minimize} \quad s \qquad\qquad s \in \mathbb{R}. \\ &\text{s.t.} \quad f_i(x) \le s, \quad i=1,\dots,m \qquad (3) \\ &\qquad Ax = b. \end{aligned}$$

- If x, s, feasible with $s<0$, then x is strictly feasible for (2)
- If optimal value $\bar{p}^*$ of (3) is positive, the problem (2) is infeasible.
- If $\bar{p}^* = 0$. and attained, then problem (3) is feasible (but not strictly).

  If $\bar{p}^* = 0$, and not attained, then (3) is infeasible.

In practice it is unlikely $\bar{p}^* = 0$.
usually $|\bar{p}^*| = \epsilon$ for small ϵ.

- use the barrier method to solve Program (3)

Suppose $x^{(0)} \in \mathrm{dom} f_1 \cap \cdots \cap \mathrm{dom} f_m$.

$Ax^{(0)} = b$.

Choose $s^{(0)} > \max\limits_{i=1,\dots,m} f_i(x^{(0)})$

$\Rightarrow f_i(x^{(0)}) < s^{(0)}$, for $i=1,\dots,m$.

Therefore $(x^{(0)}, s^{(0)})$ is a strictly feasible point for (3). $\Rightarrow$ The barrier method.

Phase I via infeasible start Newton method.

$$\begin{aligned} \text{minimize} \quad & f_0(x). \\ \text{s.t.} \quad & f_i(x) \le s, \quad i=1,\dots,m. \\ & Ax=b, \quad s=0 \in \mathbb{R}. \end{aligned}$$

$\Rightarrow$ use infeasible start Newton method to solve

$$\begin{aligned} \text{minimize} \quad & f_0(x) - \sum_{i=1}^{m} \log\left(s - f_i(x)\right). \\ \text{s.t.} \quad & Ax=b. \quad s=0. \end{aligned}$$

This can be initialized by any $x \in D$ and

$$s > \max_i f_i(x).$$

No good stopping criterion when the problem is infeasible.

Generalized inequalities.

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \preceq_{K_i} 0, \quad i=1,\dots,m \\ & Ax=b. \end{aligned}$$

- f_0 convex, $f_i : \mathbb{R}^n \to \mathbb{R}^{k_i}$ convex with respect to proper cones $K_i \in \mathbb{R}^{k_i}$.
- f_i: twice continuously differentiable.

We need a logarithm function

$\phi(x) = -\sum_{i=1}^{m} \psi(-f_i(x))$ that applies to

general proper cones.

Generalized logarithm for proper cones.

$\psi: \mathbb{R}^q \to \mathbb{R}$ is a generalized logarithm for a proper cone $K \in \mathbb{R}^q$ if

— ψ is ① concave, ② closed, ③ twice continuously

(f is closed if $S_\alpha = \{x \in \mathrm{dom}\, f : f(x) < \alpha\}$ is closed $\forall\, \alpha \in \mathbb{R}$. (A33).)

differentiable. ④ $\mathrm{dom}\, \psi = \mathrm{int}\, K$.

and ⑤ $\nabla^2 \psi(y) < 0$ for $y >_K 0$.

$f(x) = \log x$ defined on $\mathbb{R}^+$.

$f'(x) = \frac{1}{x}$. $f''(x) = -\frac{1}{x^2} < 0$.

$\log ax = \log a + \log x$. with degree $= 1$;

— There is a constant $\theta > 0$ such that

$\psi(sy) = \psi(y) + \boxed{\theta} \underline{\log s}$ for $y >_K 0$, $s > 0$. (Homogeneity)

$f(\alpha x) = \alpha^\theta \cdot f(x)$. θ is the degree.

$\Rightarrow \boxed{\log f(\alpha x)} = \underline{\theta} \log \alpha + \boxed{\log f(x)}$

Examples.

1. nonnegative orthant $K = \mathbb{R}^n_+$, For $y \in \mathbb{R}^n_+$.

$$\psi(y) = \sum_{i=1}^{n} \log(y_i).$$

$$\psi(ay) = \sum_{i=1}^{n} \log(a y_i) = \underline{n \log a} + \psi(y).$$

$\Rightarrow$ degree is n.

2. positive semidefinite cone $K = S^n_+$.

$$\psi(Y) = \log \det Y$$

$$\psi(\alpha Y) = \log \det(\alpha Y) = \log(\alpha^n \det Y).$$

$$= n \log \alpha + \log \det Y.$$

$\Rightarrow$ degree is n.

$f(X) = \underline{\log \det X}$. $\quad f'(X) = X^{-1}$.

(A.4).

Properties:

For $y \succ_{K^*} 0$,

① $\nabla \psi(y) \succ_{K^*} 0$

① $\Rightarrow$ ψ is K-increasing.

(Exercise 11.15)

② $y^T \nabla \psi(y) = \theta$.

$\frac{\partial}{\partial s}$ and set $s=1$.

$\psi(sy) = \theta \log s + \psi(y)$

Examples.

1. $\mathbb{R}^n_+$: $\psi(y) = \sum_{i=1}^n \log y_i$.

$$\nabla \psi(y) = \begin{pmatrix} 1/y_1 \\ \vdots \\ 1/y_n \end{pmatrix}. \quad y^T \nabla \psi(y) = n.$$

2. S^n_n : $\psi(Y) = \log \det Y$.

$$\nabla \psi(Y) = Y^{-1}$$

If $Y \succ 0$, then $Y^{-1} \succ 0$.

$$\mathrm{Tr}(Y \nabla \psi(Y)) = \mathrm{Tr}(Y \cdot Y^{-1}) = n.$$

Logarithmic barrier and central path.

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

$$\operatorname{dom} \phi = \{ x : f_i(x) \prec_{K_i} 0, \ i=1,\dots,m \}$$

- ψ_i is a generalized logarithm for K_i, with degree θ_i.
- ϕ is convex, twice continuously differentiable.

Central path: $\{ x^*(t) : t > 0 \}$, where $x^*(t)$ solves

$$\begin{array}{ll} \text{minimize} & t f_0(x) + \phi(x) \\ \text{s.t.} & Ax = b. \end{array}$$

Dual points on central path.

- $x^*(t)$ is optimal if there exists $w \in \mathbb{R}^p$ s.t.

$$t\nabla f_0(x) + \boxed{\sum_{i=1}^m Df_i(x)^T \nabla \psi_i(-f_i(x))} + A^T w = 0.$$

$$\left(0 = t\nabla f_0(x) + \nabla \phi(x) + A^T v\right)$$

$$\phi(x) = -\sum_i \psi(-f_i(x)).$$

$$h(x) = g(f(x)) \Rightarrow \nabla h(x) = Df(x)^T \nabla g(f(x))$$

chain rule.

$$Dh(x) = Dg(f(x)) \cdot Df(x)$$

Transpose.

- $x^*(t)$ minimizes $L(x, \lambda^*(t), v^*(t))$ (of the original optimization problem), where

$$\begin{cases} \lambda_i^*(t) = \boxed{\dfrac{1}{t} \nabla \psi_i(-f_i(x^*(t)))} \\ v^*(t) = \dfrac{w}{t}. \end{cases}$$

$$g(\lambda^*(t), v^*(t)) = f_0(x^*(t)) + \sum_{i=1}^m \underbrace{\lambda^*(t)^T \boxed{f_i(x^*(t))}}_{-\theta_i} + v^*(t)^T \underbrace{(Ax^*(t) - b)}_{=0}$$

$-y$

$$= f_0(x^*(t)) - \underbrace{\frac{1}{t}\sum_{i=1}^m \theta_i}_{\text{duality gap}}$$

For usual log.
$\theta_i = 1$. duality gap $\frac{m}{t}$.

Barrier method

given strictly feasible x, $t := t^{(0)} > 0$.
$\mu > 1$, tolerance $\epsilon > 0$.

repeat 1. Centering step.

Compute $x^*(t)$ by minimizing $tf_0 + \boxed{\phi}$
s.t. $Ax = b$.

2. Update $x := x^*(t)$.

3. Stopping criterion

. quit if $\boxed{\left(\frac{1}{t}\sum_{i=1}^{m} \theta_i < \epsilon.\right)}$

4. Increase $t := \mu t$.

Example. SDP.

minimize $c^T x$.

s.t. $F(x) = \sum_{i=1}^{n} x_i F_i + G \preceq 0$.

where $G, F_1, \dots, F_n \in \boxed{S^p}$

- log barrier function

$$\Phi(x) = -\log\det(-F(x)).$$

with degree $\underline{p}$

- central path:

$x^*(t)$ minimizes

$$t c^T x - \log\det(-F(x))$$

$$0 = \frac{\partial}{\partial x_i}(\quad) \Rightarrow \underline{t c_i - \mathrm{tr}\left(F_i F(x^*(t))^{-1}\right) = 0}.$$

$i = 1, \dots, n$.

- dual point on central path

(dual problem)

max. $\mathrm{tr}\, G Z$

s.t. $\underline{\mathrm{tr}\, F_i Z + c_i = 0}$

$Z \succeq 0$.

$$Z^*(t) = -\frac{1}{t} F(x^*(t))^{-1}$$

dual feasible.

- duality gap on central path: $\underline{\frac{p}{t}}$.