• Feasibility and Phase 1 method.

feasibility problem: find x such that
$f_i(x) < 0 \quad Ax = b$

1. basic phase 1 method

$$\begin{aligned} \min. \ (\text{over } x, s) \quad & s \\ s.t. \quad & f_i(x) - s \le 0 \\ & Ax = b \end{aligned}$$

(s is the maximum infeasibility)

if x, s feasible, with $s < 0$, then x is strictly feasible
(if s get negative, then quit)

if s^* is positive, the the original problem is infeasible
if s^* is 0 and attained, then the problem is (not strictly) feasible

push s to the left, if it stops going left a whole bunch of constraints are violated by s

2. Sum of infeasibilities phase 1 method

$$\begin{aligned} \min \quad & \mathbf{1}^T s \\ s.t. \quad & s \ge 0 \quad f_i(x) \le s_i \\ & Ax = b \end{aligned}$$

$\|s\|_1$ sparse solution

for a infeasibility problem, produce a solution that satisfies as many constraints as possible

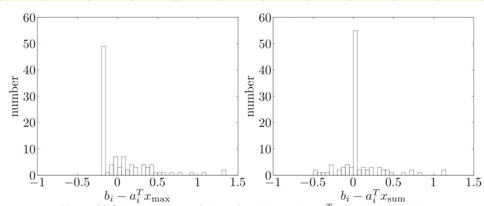
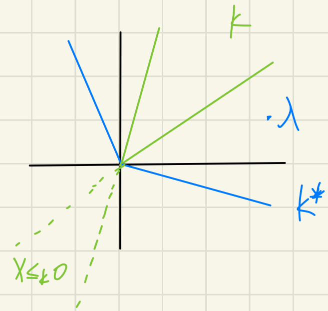


Figure 11.9 Distributions of the infeasibilities $b_i - a_i^T x$ for an infeasible set of 100 inequalities $a_i^T x \le b_i$, with 50 variables. The vector $x_{\max}$ used in the left plot was obtained by the basic phase I algorithm. It satisfies 39 of the 100 inequalities. In the right plot the vector x_{sum} was obtained by minimizing the sum of the infeasibilities. This vector satisfies 79 of the 100 inequalities.

$u = 1 + 1/\sqrt{m}$

- Generalized inequalities



$$\begin{aligned} \min.\ & f_0(x) \\ \text{s.t.}\ & f_i(x) \leq_{K_i} 0 \\ & Ax = b \end{aligned}$$

(SDP SOCP)

$$L(x, \lambda, \nu) = f_0(x) + \lambda_i^T f_i(x) + \nu^T(Ax-b)$$

$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu)$$

if $\tilde{x}$ is feasible and $\lambda_i \geq_{K_i^*} 0$

$$f(\tilde{x}) \geq L(\tilde{x}, \lambda, \nu) \geq \inf_{x \in D} L(x, \lambda, \nu) = g(\lambda, \nu)$$

$$\begin{aligned} \max.\ & g(\lambda, \nu) \\ \text{s.t.}\ & \lambda_i \geq_{K_i^*} 0 \end{aligned}$$

if strong duality holds

$$\begin{aligned} f_0(x^*) = g(\lambda^*, \nu^*) &= \inf_{x \in D} \left[f_0(x) + \sum_i \lambda_i^{*T} f_i(x) + \sum_i \nu_i h_i(x) \right] \\ &\leq f_0(x^*) + \sum_i \lambda_i^{*T} f_i(x^*) + \sum_i \nu_i^* h_i(x) \\ &\leq f_0(x^*) \end{aligned}$$

KKT conditions

$$Ax^* = b$$

$$f_i(x^*) \leq_{K_i} 0$$

$$\lambda_i^* \geq_{K_i^*} 0$$

$$\nabla f_0(x^*) + \sum_i Df_i(x^*)^T \lambda_i^* + A^T \nu^* = 0$$

$$\lambda_i^{*T} f_i(x^*) = 0$$

• Generalized logarithm for proper cone

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

- $\mathrm{dom}\, \psi = \mathrm{int}\, K$ and $\nabla^2 \psi(y) < 0$ for $y \succ_K 0$ — e.g. for $x \le 0$ $\mathrm{dom} \log(-x) = \{x < 0\}$
- $\psi(sy) = \psi(y) + \theta \cdot \log s$ for $y \succ_K 0$ $s > 0$ (θ is the degree of ψ)

eg. non-negative orthant $K = \mathbb{R}^n_+$: $\psi(y) = \sum_i \log y_i$, degree $\theta = n$

eg. positive semi-definite cone $K = S^n_+$

$$\psi(Y) = \log \det(Y) \qquad \theta = n$$

eg. Second-order cone $K = \{ y \in \mathbb{R}^{n+1} \mid (y_1^2 + \cdots + y_n^2)^{1/2} \le y_{n+1} \}$

$$\psi(y) = \log(y_{n+1}^2 - y_1^2 \cdots - y_n^2) \qquad \theta = 2$$

Generalized logarithm is K-increasing

if $y \succ_K 0$, then $\nabla \psi(y) \succeq_{K^*} 0$

if $\nabla\psi(y) \nsucceq_{K^*} 0$, $s > 0$

then there exists $w \succ_K 0$ s.t. $w^T \nabla\psi(y) \le 0$

$\psi(sw) \le \psi(y) + \nabla\psi(y)^T(sw - y)$

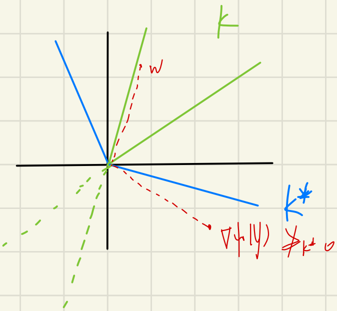
$= \psi(y) + s \cdot \nabla\psi(y)^T w - y^T \nabla\psi(y)$

$\le \psi(y) - y^T \nabla\psi(y)$

$= \psi(y) - \theta$

$\psi(sw)$ is bounded above

but $\psi(sw) = \psi(w) + \theta \cdot \log s$ is unbounded



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

$\nabla\psi(sy)^T y = \frac{\theta}{s}$

$(sy)^T \nabla\psi(sy) = \theta$

$$y^T \nabla\psi(y) = \theta$$

eg. positive semi-def cone S^n_+ : $\psi(Y) = \log\det(Y)$

$\nabla\psi(Y) = Y^{-1}$ $\qquad tr(Y^T Y^{-1}) = n$

eg. second-order cone $K = \{ y \in \mathbb{R}^{n+1} \mid (y_1^2 + \cdots y_n^2)^{1/2} \le y_{n+1} \}$

$$\nabla\psi(y) = \frac{2}{y_{n+1}^2 - y_1^2 \cdots - y_n^2} \cdot \begin{bmatrix} -y_1 \\ \vdots \\ -y_n \\ y_{n+1} \end{bmatrix}$$

- Logarithmic barrier and central path

logarithmic barrier fn : $f_1(x) \preceq_{K_1} 0 \quad \cdots \quad f_m(x) \preceq_{K_m} 0$

$$\phi(x) = -\sum_i \phi_i(-\log f_i(x)) \qquad \text{dom } \phi_i = \{x \mid f_i(x) \prec_{K_i} 0\}$$

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

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

- Dual points on central path

$$\mathcal{L}(x,v) = t \cdot f_0(x) + \sum_i \phi_i(-f_i(x)) + v^T(Ax-b) = 0$$

optimality condition $\begin{cases} \nabla_x \mathcal{L}(x,v) = t \cdot \nabla f_0(x) + \sum_i Df_i(x)^T \cdot \nabla\phi_i(-f_i(x)) + A^T v = 0 \\ Ax-b=0 \end{cases}$

$$\nabla f_0(x^*(t)) + \sum_i Df(x)^T \cdot \underbrace{\nabla\phi_i(-f_i(x^*(t)))/t}_{\lambda_i} + A^T \underbrace{(v/t)}_{v^*(t)} = 0$$

$x^*(t)$ minimizes the Lagrangian

$$\mathcal{L}(x, \lambda^*(t), v^*(t)) = f_0(x) + \sum_i f_i(x^*(t))^T \lambda_i^*(t) + v^*(t)^T(Ax-b)$$

where $\lambda_i^*(t) = \nabla\phi_i(-f_i(x^*(t)))/t \quad v^*(t) = v/t$ dual feasible

if $y \succ_K 0$ then $\nabla\phi(y) \succeq_{K^*} 0$

$$\begin{aligned} \min\ & f_0(x) \\ \text{s.t.}\ & f_i(x) \preceq_{K_i} 0 \\ & Ax=b \end{aligned} \qquad \begin{aligned} \mathcal{L}(x,\lambda,v) &= f_0(x) + \sum_i \lambda_i^T f_i(x) + v^T(Ax-b) \\ g(\lambda,v) &= \inf_x \mathcal{L}(x,\lambda,v) \end{aligned}$$

$$\left.\begin{aligned} \min.\ & t f_0(x) + \phi(x) \\ \text{s.t.}\ & Ax=b \end{aligned}\right\}$$ solving this problem gives a primal-dual pair thus a lower bound of the original problem

$$\begin{aligned} p^* &\geq g(\lambda^*(t), v^*(t)) \\ &= \mathcal{L}(x^*(t), \lambda^*(t), v^*) \\ &= f_0(x^*(t)) + \sum_i f_i(x^*(t))^T \nabla\phi_i(-f_i(x^*(t)))/t + v^*(t)^T(Ax^*(t)-b) \\ &= f_0(x^*(t)) - \sum_i \theta_i/t \qquad (\theta_i \text{ is the degree of } \phi_i) \end{aligned}$$

eg. semidefinite programming

$$\min.\ c^T x$$
$$\text{s.t. } F(x) = \sum_i F_i \cdot x_i + G \preceq 0 \qquad (F_i, G \in S^p)$$

logarithmic barrier : $\phi(x) = -\log\det(-F(x)) = \log\det(-F(x)^{-1})$

$$\min.\ t \cdot c^T x + \log\det(-F(x)^{-1}) \quad = f_0(x)$$
$$\text{s.t. } Ax = b$$

$$\mathcal{L}(x,v) = f_0(x) + v^T(Ax-b)$$

$\left\{\begin{array}{l} \nabla\mathcal{L}(x,v) = 0 \\ Ax = b \end{array}\right\}$ optimality condition $\quad \left\{\begin{array}{l} \nabla\mathcal{L}(x+\Delta x, v) \\ A\Delta x = 0 \end{array}\right\}$ newton step

$$\frac{\partial f_0}{\partial x_i} = t \cdot c_i + tr(F_i^T F(x)^{-1}) = t \cdot c_i - tr(F_i\, F(x)^{-1})$$

$$\nabla_{x_i} f_0(x+\Delta x) = t\cdot c_i - tr[F_i\, F(x+\Delta x)^{-1}]$$
$$= t\cdot c_i - tr[F_i\, (F(x) + \sum_j F_j \Delta x_j)^{-1}]$$
$$\approx t\cdot c_i - tr[F_i\, (F(x)^{-1} - F(x)^{-1}(\sum_j F_j \Delta x_j) F(x)^{-1})]$$
$$= t\cdot c_i - tr[F_i\, F(x)^{-1}] + tr[F_i\, F(x)^{-1} (\sum_j F_j \Delta x_j)\, F(x)^{-1}]$$
$$= t\cdot c_i - tr[F_i F(x)^{-1}] + \sum_j \Delta x_j \cdot tr[F_i\, F(x)^{-1} F_j F(x)^{-1}] := 0$$

$$\frac{\partial \mathcal{L}}{\partial x_i} = t\cdot c_i - tr[F_i F(x)^{-1}] + \sum_j \Delta x_j \cdot tr[F_i F(x)^{-1} F_j F(x)^{-1}] + (A^T v)_i = 0$$

newton step $\left\{\begin{array}{l} \sum_j \Delta x_j \cdot tr[F_i\, F(x)^{-1} F_j F(x)^{-1}] + (A^T v)_i = tr[F_i F(x)^{-1}] - t\cdot c_i \\ A\Delta x = 0 \end{array}\right.$

KKT system $\left\{ \begin{bmatrix} TR & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ v \end{bmatrix} = \begin{bmatrix} F_i \cdot F(x)^{-1} \end{bmatrix} - t \begin{bmatrix} c \end{bmatrix} \right.$

$$\mathcal{L} = tc^T x + \log\det(-F(x)) + v^T(Ax-b)$$
$$\nabla_{x_i}\mathcal{L} = t\cdot c_i - tr(F_i F(x)^{-1}) + (A^T v)_i$$

$$t\cdot c_i - tr(F_i\, F(x^*(t))^{-1}) + (A^T v)_i = 0$$
$$c_i - tr(F_i\, F(x^*(t))^{-1}/t) + (A^T v/t)_i = 0$$

$\to x^*(t)$ minimize the lagrangian

$$\mathcal{L}(x, \lambda^*(t), v^*(t))$$
$$= c^T x + tr(F(x)\lambda^*(t)) + v^*(Ax-b)$$

$\lambda^*(t) = -(1/t)\cdot F(x^*(t))^{-1}$ dual point $\in S^p_{++}$

$$p^* \geq g(\lambda^*(t), v^*(t)) = \mathcal{L}(x^*(t), \lambda^*(t), v^*(t))$$
$$= c^T x^*(t) - 1/t\ tr(F(x^*(t))\cdot F(x^*(t))^{-1}) + v^T(Ax^*(t)-b)$$
$$= c^T x^*(t) - p/t \quad \text{duality gap}$$