6.3

Formulate following approximation problems as LPs, QPs, SOCPs or SDPs

soln

△ first recall LP form

$$\begin{aligned} \text{minimize} \quad & c^T x + d \\ \text{subject to} \quad & Gx \preceq h \\ & Ax = b \end{aligned}$$

△ first recall QP form

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2} x^T P x + q^T x + r \\ \text{subject to} \quad & Gx \preceq h \\ & Ax = b \end{aligned}$$

△ first recall SOCP form

$$\begin{aligned} \text{minimize} \quad & f^T x \\ \text{subject to} \quad & \|A_i x + b_i\|_2 \leq c_i^T x + d_i \\ & Fx = g \end{aligned}$$

△ first recall SDP form

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & x_1 F_1 + x_2 F_2 + \cdots + x_n F_n + G \preceq 0 \\ & Ax = b \end{aligned}$$

(a) Deadzone-linear penalty approximation:

△ $\underset{x}{\text{minimize}} \sum_{i=1}^{m} \phi(a_i^T x - b_i)$

where $\phi(u) = \begin{cases} 0 & |u| \leq a \\ |u| - a & |u| > a \end{cases}$

$a > 0$

$\Rightarrow \underset{x,t}{\text{minimize}} \sum_{i=1}^{m} t_i$ (through epigraph problem form)

subject to $\phi(a_i^T x - b_i) \leq t_i$

$\phi(a_i^T x - b_i) = \begin{cases} 0 & |a_i^T x - b_i| \leq a \\ |a_i^T x - b_i| - a & |a_i^T x - b_i| > a \end{cases}$

$\Rightarrow \underset{x,t}{\text{minimize}} \quad 1^T t \qquad t \in R^m$

subject to $0 \preceq t$

$|Ax - b| - a1 \preceq t$

$\Rightarrow \underset{x,t}{\text{minimize}} \quad 1^T t \qquad t \in R^m$

subject to $0 \preceq t, \; -t - a1 \preceq Ax - b \preceq t + a1$ △ which is an LP ☆

6.3 (cont'd)

soln (cont'd)

(b) Log-barrier penalty approximation : $\forall i = 1, \dots, m$

Δ $\underset{x}{\text{minimize}} \sum_{i=1}^{m} \phi(a_i^T x - b_i)$

where $\phi(u) = \begin{cases} -a^2 \log(1-(\frac{u}{a})^2) & |u| < a \\ \infty & |u| \geq a \end{cases}$

$a > 0$

$\Rightarrow \underset{x}{\text{minimize}} \sum_{i=1}^{m} -a^2 \log(1 - (\frac{a_i^T x - b_i}{a})^2)$

subject to $|a_i^T x - b_i| < a$

$\Rightarrow \underset{x,y}{\text{minimize}} \sum_{i=1}^{m} -a^2 \log(1-(\frac{y_i}{a})^2)$

subject to $|y_i| < a$

$y_i = a_i^T x - b_i$

$\sum_{i=1}^{m} -a^2 \log(1-(\frac{y_i}{a})^2)$

$= -a^2 \log(1-(\frac{y_1}{a})^2) + -a^2 \log(1-(\frac{y_2}{a})^2) + \cdots + -a^2 \log(1-(\frac{y_m}{a})^2)$

$= -a^2 \log(\prod_{i=1}^{m}(1-(\frac{y_i}{a})^2))$

$\Rightarrow \underset{x,y}{\text{minimize}} -a^2 \log(\prod_{i=1}^{m}(1-(\frac{y_i}{a})^2))$

subject to $|y_i| < a$

$y_i = a_i^T x - b_i$

$\Rightarrow \underset{x,y}{\text{maxiximize}} \prod_{i=1}^{m}(1-(\frac{y_i}{a})^2)$

subject to $|y_i| < a$

$y_i = a_i^T x - b_i$

here we then try to induce hyperbolic constraints & epigraph form

$\left\|\begin{bmatrix} 2x \\ y-z \end{bmatrix}\right\|_2 \leq y+z, \ y \geq 0, z \geq 0$

$\Updownarrow$

$\left\|\begin{bmatrix} 2x \\ y-z \end{bmatrix}\right\|_2^2 \leq (y+z)^2, \ y \geq 0, z \geq 0$

$\Updownarrow$

$4x^Tx + (y-z)^2 \leq y^2 + z^2 + 2yz, \ y \geq 0, z \geq 0$

$\Updownarrow$

$x^T x \leq yz, \ y \geq 0, z \geq 0$

$\Rightarrow \underset{x,y,t}{\text{maximize}} \prod_{i=1}^{m} t_i^2$

subject to $|y_i| < a$

$y_i = a_i^T x - b_i$

$1-(\frac{y_i}{a})^2 \geq t_i^2$

$\Rightarrow \underset{x,y,t}{\text{maximize}} \prod_{i=1}^{m} t_i^2$

subject to $-a < y_i < a$

$y_i = a_i^T x - b_i$

$(1+\frac{y_i}{a})(1-\frac{y_i}{a}) \geq t_i^2$

$\Rightarrow \underset{x,y,t}{\text{maximize}} t_1^2 t_2^2 t_3^2 t_4^2$

subject to $-a < y_i < a$

$y_i = a_i^T x - b_i$

$(1+\frac{y_i}{a})(1-\frac{y_i}{a}) \geq t_i^2$

we assume, without loss of generality, take $m = 2^K = 4, K \in N$ $\forall i = 1,2,3,4$ as example

$\Rightarrow \underset{x,y,t,v}{\text{maximize}} (V_1^2)^2 (V_2^2)^2$

subject to $-a < y_i < a$

$y_i = a_i^T x - b_i$

$(1+\frac{y_i}{a})(1-\frac{y_i}{a}) \geq t_i^2$ $\quad \forall i = 1,2,3,4$

$t_1 t_2 \geq V_1^2 \quad t_3 t_4 \geq V_2^2$

6.3 (cont'd)
sol'n (cont'd)
(b) (cont'd)

$$\Rightarrow \underset{x,y,t,V}{\text{maximize}} \quad V_1^2 V_2^2$$
$$\text{subject to} \quad -a < y_i < a$$
$$y_i = a_i^T x - b_i$$
$$\left(1+\frac{y_i}{a}\right)\left(1-\frac{y_i}{a}\right) \geq t_i^2 \quad \forall i=1,2,3,4$$
$$t_1 t_2 \geq V_1^2$$
$$t_3 t_4 \geq V_2^2$$

$$\Rightarrow \underset{x,y,t,V,W}{\text{maximize}} \quad W^2 \quad (\Leftrightarrow \text{maximize } (W^2)^2)$$
$$-a < y_i < a$$
$$y_i = a_i^T x - b_i$$
$$\left(1+\frac{y_i}{a}\right)\left(1-\frac{y_i}{a}\right) \geq t_i^2 \quad \forall i=1,2,3,4$$
$$t_1 t_2 \geq V_1^2$$
$$t_3 t_4 \geq V_2^2$$
$$V_1 V_2 \geq W^2$$

$$\Rightarrow \underset{x,y,t,V,W,z}{\text{maximize}} \quad z$$
$$\left.\begin{array}{l} -a < y_i < a \\ y_i = a_i^T x - b_i \\ \left(1+\frac{y_i}{a}\right)\left(1-\frac{y_i}{a}\right) \geq t_i^2 \end{array}\right\} \quad \forall i=1,2,3,4$$
$$t_1 t_2 \geq V_1^2$$
$$t_3 t_4 \geq V_2^2$$
$$W^2 \geq z$$

$$\Rightarrow \underset{x,y,t,V,W,z}{\text{minimize}} \quad -z$$
$$\|0x+0\|_2 \leq a - I_i^T x \quad i=1,2,3,4$$

identity matrix

$$\left\|\begin{bmatrix} 2t_i \\ \frac{2y_i}{a} \end{bmatrix}\right\|_2 \leq 2 \quad i=1,2,3,4$$
$$\left\|\begin{bmatrix} 2V_1 \\ t_1 - t_2 \end{bmatrix}\right\|_2 \leq t_1 + t_2$$
$$\left\|\begin{bmatrix} 2V_2 \\ t_3 - t_4 \end{bmatrix}\right\|_2 \leq t_3 + t_4$$
$$\left\|\begin{bmatrix} 2W \\ z-1 \end{bmatrix}\right\|_2 \leq z+1$$
$$y = Ax - b$$

$\triangle$ which is re-written in SOCP

✲

6.3 (cont'd)

sol'n (cont'd)

(c) Huber penalty approximation

△ minimize $\sum_{i=1}^{m} \phi(a_i^T x - b_i)$

where $\phi(u) = \begin{cases} u^2 & |u| \le M \\ M(2|u| - M) & |u| > M \end{cases}$

$M > 0$

⇒ as shown in HW3, 4.5, the above is equivalent to

$$\begin{aligned} \underset{x,u,v}{\text{minimize}} \quad & \sum_{i=1}^{m} (u_i^2 + 2Mv_i) \\ \text{subject to} \quad & -u - v \preceq Ax - b \preceq u + v \\ & 0 \preceq u \preceq M\mathbf{1} \\ & v \succeq 0 \end{aligned}$$

△ which is expressed in QP ☆

(d) Log-Chebyshev approximation

△ minimize $\max_{i=1,\dots,m} |\log(a_i^T x) - \log b_i|$ $b \succ 0$ assumed.

$$\Leftrightarrow \quad \begin{aligned} \underset{x,t}{\text{minimize}} \quad & t \\ \text{subject to} \quad & \frac{1}{t} \le a_i^T x / b_i \le t, \quad i = 1, \dots, m. \end{aligned}$$

here indure hyperbolic constraints

$$\Rightarrow \quad \begin{aligned} \underset{x,t}{\text{minimize}} \quad & t \\ \text{subject to} \quad & \left.\begin{aligned} \frac{1}{t} \le \frac{a_i^T x}{b_i} \\ \frac{a_i^T x}{b_i} \le t \end{aligned}\right\} \quad i = 1, \dots, m \end{aligned}$$

$$\Rightarrow \quad \begin{aligned} \underset{x,t}{\text{minimize}} \quad & t \\ & b_i \le t\, a_i^T x \\ & t \ge 0 \\ & a_i^T x \le b_i t \end{aligned}$$

$$\Rightarrow \quad \begin{aligned} \underset{x,t}{\text{minimize}} \quad & t \\ & \left\| \begin{bmatrix} 2\sqrt{b_i} \\ t - a_i^T x \end{bmatrix} \right\|_2 < t + a_i^T x \\ & \|0\|_2 < -t \\ & \|0\|_2 \le b_i t - a_i^T x \end{aligned}$$

△ which is an SOCP ☆

6.3 (cont'd)

soln (cont'd)

(e) Minimizing sum of the largest k residuals

△ minimize $\sum_{i=1}^{k} |r|_{[i]}$

subject to $r = Ax - b$

where $|r|_{[1]} \geq |r|_{[2]} \geq \cdots \geq |r|_{[m]}$

$\Rightarrow \underset{t,u}{\text{minimize}} \quad \sum_{i=1}^{m} t_i + \sum_{i=1}^{k} u$

subject to $-t_i \leq a_i^T x - b_i \leq t_i$

$|a_1^T x - b_1| \leq u$

$|a_k^T x - b_k| \leq u$

$\Rightarrow \underset{t,u}{\text{minimize}} \quad 1^T t + ku$

subject to $-t - u1 \preceq Ax - b \preceq t + u1$

$t \succeq 0, \; u \succeq 0$

△ which is a LP ☆

6.8 formulate below robust approximation problems as LP, QP, SOCP, or SDP

for each, derive l_1-, l_2- & l_∞- norm.

soln

(a) △ minimize $\sum_{i=1}^{k} p_i \overset{\text{got } i=1,2,3\ldots k \; \langle A_i, p_i \rangle}{\|A_i x - b\|_1}$, $p \succeq 0$, $1^T p = 1$

$\Rightarrow$ minimize $\sum_{i=1}^{k} p_i 1^T \overset{R^m}{t_i}$

subject to $-t_i \preceq A_i x - b \preceq \overset{R^m}{t_i}$ $\quad i=1,\ldots k.$

$t_i \geq 0$

△ which is an LP (l_1-norm) ☆

△ minimize $\sum_{i=1}^{k} p_i \|A_i x - b\|_2 \quad p \succeq 0, \; 1^T p = 1$

$\Rightarrow$ minimize $\sum_{i=1}^{k} p_i \overset{R}{t_i}$

subject to $\|A_i x - b\|_2 \leq t_i \quad i=1,\ldots k$

△ whih is an SOCP (l_2-norm) ☆

6.8 (cont'd)

sol^n (cont'd)

(a) (cont'd)

Δ minimize $\sum_{i=1}^{k} p_i \|A_i x - b\|_\infty$ $\quad p \succeq 0,\ 1^T p = 1$

$\Rightarrow$ minimize $\sum_{i=1}^{k} p_i \overset{\mathbb{R}}{t_i}$

subject to $-t_i 1 \preceq \overset{\mathbb{R}^m}{A_i x - b} \preceq t_i 1 \quad i=1,\dots k$

Δ which is an LP $\times$

(b)

minimize $\sup_{A \in \mathcal{A}} \|Ax - b\|$

where $\mathcal{A} = \{ A \in \mathbb{R}^{m\times n} \mid l_{ij} \le a_{ij} \le u_{ij},\ i=1,\dots,m,\ j=1,\dots n \}$

l_1-norm:

Δ minimize $\sup_{A\in\mathcal{A}} \|Ax-b\|_1$

$\Rightarrow$ minimize $\sup_{A\in\mathcal{A}} \sum_{i=1}^{m} | a_i^T x - b_i |$

$$\sup_{l_{ij}\le a_{ij}\le u_{ij}} |a_i^T x - b_i|$$

$$= \sup_{l_{ij}\le a_{ij}\le u_{ij}} \left|\sum_{j=1}^{n} a_{ij} x_j - b_i\right|$$

$$= |\bar{a}_i^T x - b_i| + \sum_{j=1}^{n} v_{ij} |x_j|$$

$\boxed{\bar{a}_{ij} = \frac{(l_{ij}+u_{ij})}{2}}$ mid-point

$\boxed{v_{ij} = \frac{(u_{ij}-l_{ij})}{2}}$ positive direction (as far as possible within the $l_{ij} \le a_{ij} \le u_{ij}$ interval.

$\Rightarrow$ minimize $\sum_{i=1}^{m} \left(|\bar{a}_i^T x - b_i| + \sum_{j=1}^{n} v_{ij} |x_j| \right)$

$\Rightarrow$ minimize $1^T (t + Vu)$

subject to
$$-t_i \le \bar{a}_i^T x - b_i \le t_i \quad i=1,\dots m$$
$$-u_i \le x_i \le u_i \quad i=1,\dots n$$
$$t_i, u_i \ge 0$$

Δ which is a LP $\times$

l_2-norm:

Δ minimize $\sup_{A\in\mathcal{A}} \|Ax-b\|_2$

$\Rightarrow$ minimize $\sup_{A\in\mathcal{A}} \left(\sum_{i=1}^{m} |a_i^T x - b_i|^2 \right)^{\cancel{\frac{1}{2}}}$

$\Rightarrow$ minimize $\sum_{i=1}^{m} \left(|a_i^T x - b_i| + \sum_{j=1}^{n} v_{ij} |x_j| \right)^2$

$\Rightarrow$ minimize v

$$-t_i \le \bar{a}_i^T x - b_i \le t_i \quad i=1,\dots m \qquad t_i \ge 0$$
$$-u_i \le x_i \le u_i \quad i=1,\dots n \qquad u_i \ge 0$$
$$\| t + Vu \|_2 \le t$$

Δ which is a SOCP $\times$

l_∞-norm :

Δ minimize $\sup_{A \in \mathcal{A}} \|Ax-b\|_\infty$

$\Rightarrow$ minimize $\max_{i=1,\dots,m} \left(|\bar{a}_i^T x - b_i| + \sum_{j=1}^{n} V_{ij} |x_j| \right)$

$\Rightarrow$ minimize V

$$\begin{aligned} -t_i &\le \bar{a}_i^T x - b_i \le t_i && i=1,\dots m && t_i \ge 0 \\ -u_i &\le \quad x_i \le u_i && i=1,\dots n && u_i \ge 0 \\ -V &\le (t+Vu)_i \le V && && V \ge 0 \end{aligned}$$

Δ which is a LP $\divideontimes$

(c) Δ minimize $\sup_{A \in \mathcal{A}} \|Ax-b\|$

where $\mathcal{A} = \{ [a_1 \dots a_m]^T \mid \underbrace{C_i}_{R^{p_i \times n}} a_i \preceq \underbrace{d_i}_{R^{p_i}},\ i=1,\dots,m \}$

Δ $\sup_{a_i \in P_i} |a_i^T x - b_i| = \max \left\{ \sup_{a_i \in P_i} (a_i^T x) - b_i,\ \sup_{a_i \in P_i} (-a_i^T x) + b_i \right\}$

$$\left.\begin{aligned} \Rightarrow \sup_{a_i \in P_i} a_i^T x &= \inf \{ d_i^T v \mid C_i^T v = x,\ v \succeq 0 \} \\ \sup_{a_i \in P_i} (-a_i^T x) &= \inf \{ d_i^T w \mid C_i^T w = -x,\ w \succeq 0 \} \end{aligned}\right\} \rightarrow \text{LP duality}$$

maximize $c^T x$
s.t. $Ax \le b$
$x \ge 0$

$\Updownarrow$

minimize $b^T y$
s.t. $A^T y \ge c$
$y \ge 0$

$\Rightarrow$ original problem becomes :

minimize $\|t\|$

subject to

$$\left.\begin{aligned} &x = C_i^T v_i \\ &x = C_i^T w_i \\ &d_i^T v_i \le t_i \\ &d_i^T w_i \le t_i \\ &v_i \succeq 0 \\ &w_i \succeq 0 \end{aligned}\right\} \quad i=1,\dots m$$

6.8 (cont'd)

soln (cont'd)

(c) (cont'd)

△ l_1-norm:

$$\text{minimize} \quad \sup_{A \in \mathcal{A}} \|Ax - b\|_1$$

$\Rightarrow$

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{m} |t_i| \\ \text{subject to} & x = C_i^T v_i \\ & x = C_i^T w_i \\ & d_i^T v_i \le t_i \\ & d_i^T w_i \le t_i \\ & v_i \succeq 0 \\ & w_i \succeq 0 \end{array}$$

$\Rightarrow$

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{m} u_i \\ \text{subject to} & x = C_i^T v_i \\ & x = C_i^T w_i \\ & d_i^T v_i \le t_i \\ & d_i^T w_i \le t_i \\ & v_i \succeq 0 \\ & w_i \succeq 0 \\ & -u_i \le t_i \le u_i \\ & u_i \ge 0 \end{array}$$

△ ∴ LP

△ l_2-norm

$$\text{minimize} \quad \sup_{A \in \mathcal{A}} \|Ax - b\|_2$$

$\Rightarrow$

$$\begin{array}{ll} \text{minimize} & \|t\|_2 \\ \text{subject to} & x = C_i^T v_i \\ & x = C_i^T w_i \\ & d_i^T v_i \le t_i \\ & d_i^T w_i \le t_i \\ & v_i \succeq 0 \\ & w_i \succeq 0 \\ & \|t\|_2 \le u \end{array}$$

△ ∴ SOCP

△ l_∞-norm

$$\text{minimize} \quad \sup_{A \in \mathcal{A}} \|Ax - b\|_\infty$$

$\Rightarrow$

$$\begin{array}{ll} \text{minimize} & \|t\|_\infty \\ \text{subject to} & x = C_i^T v_i \\ & x = C_i^T w_i \\ & d_i^T v_i \le t_i \\ & d_i^T w_i \le t_i \\ & v_i \succeq 0 \\ & w_i \succeq 0 \end{array}$$

$\Rightarrow$

$$\begin{array}{ll} \text{minimize} & u \\ \text{subject to} & x = C_i^T v_i \\ & x = C_i^T w_i \\ & d_i^T v_i \le t_i \\ & d_i^T w_i \le t_i \\ & v_i \succeq 0 \\ & w_i \succeq 0 \\ & -u\mathbf{1} \preceq t \preceq u\mathbf{1} \\ & u \succeq 0 \end{array}$$

△ ∴ LP

16 Maximum volume rectangle inside polyhedron :

find $R = \{x \in R^n \mid l \preceq x \preceq u\}$ of maximum volume

which $\subseteq P = \{x \mid Ax \preceq b\}$

with variables $l, u \in R^n$

<u>sol^n</u>

- △ for $R \subseteq P$, we need to confine all vertices within $P = \{x \mid Ax \preceq b\}$
- △ according to problem description, there can be no exponential no. of constraints.
- △ we can confine the rectangle, $R \subseteq P$ by following operation

$$\sum_{j=1}^{n} a_{ij} x_j = \sum_{j \in V_i^+} a_{ij} x_j + \sum_{j \in V_i^-} a_{ij} x_j \text{ , where } V_i^+ = \{j \mid a_{ij} \geq 0\}, V_i^- = \{j \mid a_{ij} \leq 0\}$$

$$= \sum_{j \in V_i^+} a_{ij}^+ x_j - \sum_{j \in V_i^-} a_{ij}^- x_j \leq \boxed{\sum_{j \in V_i^+} a_{ij}^+ u_j - \sum_{j \in V_i^-} a_{ij}^- l_j}$$

$\therefore$ $R \subseteq P$ i.f.f.

$$\sum_{i=1}^{n} (a_{ij}^+ u_j - a_{ij}^- l_j) \leq b_i \text{ , } i = 1, \dots, m$$

$\therefore$ problem could be formulated as

$$\begin{array}{ll} \underset{u,l}{\text{maximize}} & \prod_{i=1}^{n} (u_i - l_i) \\ \text{subject to} & \sum_{i=1}^{n} (a_{ij}^+ u_j - a_{ij}^- l_j) \leq b_i \text{ , } i = 1, \dots m \\ & u_j - l_j \geq 0 \end{array}$$

$$\Rightarrow \begin{array}{ll} \underset{u,l}{\text{maximize}} & \sum_{i=1}^{n} \log (u_i - l_i) \\ \text{subject to} & \sum_{i=1}^{n} (a_{ij}^+ u_j - a_{ij}^- l_j) \leq b_i \text{ , } i = 1, \dots m \\ & u_j - l_j \geq 0 \end{array}$$

✡

9.5 Backtracking line search

f is strongly convex w/ $mI \preceq \nabla^2 f(x) \preceq MI$

- show:

backtracking stopping condition holds for

$$0 < t \le -\frac{\nabla f(x)^T \Delta x}{M \|\Delta x\|_2^2} \quad — ⓐ$$

- give an upper bound on number of backtracking iterations

soln

Δ recall Backtracking line search

1 given Δx for f @ $x \in \operatorname{dom} f$

$\alpha \in (0, 0.5)$

$\beta \in (0, 1)$.

2 $t := 1$

3 while $f(x + t\Delta x) > f(x) + \alpha t \nabla f(x)^T \Delta x$ — ①

$t := \beta t$

Δ strong convexity implies

$$f(y) \le f(x) + \nabla f(x)^T (y - x) + \frac{M}{2} \|y - x\|_2^2$$

$$\Rightarrow f(x + t\Delta x) \le f(x) + \nabla f(x)^T (x + t\Delta x - x) + \frac{M}{2} \|x + t\Delta x - x\|_2^2$$

$$\Rightarrow f(x + t\Delta x) \le f(x) + t\nabla f(x)^T \Delta x + \frac{M}{2} t^2 \|\Delta x\|_2^2 \quad — ②$$

Δ from ①

stopping condition occurs:

$$f(x + t\Delta x) \le f(x) + \alpha t \nabla f(x)^T \Delta x \quad — ③$$

$\therefore$ from ②, ③

$$\cancel{f(x)} + \alpha t \nabla f(x)^T \Delta x - \cancel{f(x)} - t \nabla f(x)^T \Delta x - \frac{M}{2} t^2 \|\Delta x\|_2^2 \ge 0$$

$$\Rightarrow (\alpha - 1) \cancel{t} \nabla f(x)^T \Delta x - \frac{M}{2} t^{\cancel{2}} \|\Delta x\|_2^2 \ge 0$$

$$\Rightarrow (\alpha - 1) \nabla f(x)^T \Delta x - \frac{M}{2} t \|\Delta x\|_2^2 \ge 0$$

$$\Rightarrow \frac{(\alpha - 1) \nabla f(x)^T \Delta x}{\frac{M}{2} \|\Delta x\|_2^2} \ge t$$

$\triangle$ from above, when reach stopping condition

$$t \le \frac{2(\alpha-1)\nabla f(x)^T \Delta x}{M \|\Delta x\|_2^2}$$

$$\Rightarrow \quad t_0 = \frac{2(\alpha-1)\nabla f(x)^T \Delta x}{M \|\Delta x\|_2^2}$$

$\triangle$ to get ②, we have

$0 < t < t_0$ for backtracking line search

$$t_0 = \frac{2(\alpha-1)\nabla f(x)^T \Delta x}{M \|\Delta x\|_2^2} \ge \frac{-\nabla f(x)^T \Delta x}{M \|\Delta x\|_2^2}$$

$$\therefore \quad 0 < t \le \frac{-\nabla f(x)^T \Delta x}{M \|\Delta x\|_2^2}$$

$\triangle$ t initially $t := 1$

$$\beta^k \le \frac{-\nabla f(x)^T \Delta x}{M \|\Delta x\|_2^2}$$

$$\Rightarrow \quad \log \beta^k \le \log \left(\frac{-\nabla f(x)^T \Delta x}{M \|\Delta x\|_2^2} \right)$$

$$\Rightarrow \quad k \le \frac{\log \left(\frac{-\nabla f(x)^T \Delta x}{M \|\Delta x\|_2^2} \right)}{\log \beta}$$

9.9 Newton decrement . show that $\lambda(x)$

$$\lambda(x) = \sup_{v^T \nabla^2 f(x) v = 1} \left(-v^T \nabla f(x) \right) = \sup \frac{-v^T \nabla f(x)}{(v^T \nabla^2 f(x) v)^{1/2}}$$

<u>sol^n</u>

△ recall Newton decrement :

$$\lambda(x) = \left(\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x) \right)^{1/2}$$

△ $\sup_{v^T \nabla^2 f(x) v = 1} \left(-v^T \nabla f(x) \right)$ $\qquad u = \nabla^2 f(x)^{\frac{1}{2}} v$

$\qquad v = \nabla^2 f(x)^{-\frac{1}{2}} u$

$$\Rightarrow \sup_{\|u\|_2^2 = 1} \left(-\left(\nabla^2 f(x)^{-\frac{1}{2}} u \right)^T \nabla f(x) \right)$$

$$\Rightarrow \sup_{\|u\|_2 = 1} \left(-u^T \nabla^2 f(x)^{-\frac{1}{2}} \nabla f(x) \right)$$

$$\Rightarrow \sup_{\|u\|_2 = 1} \left(\nabla f(x)^T u^T \nabla^2 f(x)^{-1} u \nabla f(x) \right)^{\frac{1}{2}}$$

$$\Rightarrow \left(\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x) \right)^{\frac{1}{2}} = \lambda(x)$$

△ which the supremum of $-v^T \nabla f(x)$, $v^T \nabla^2 f(x) v = 1$ is $\lambda(x)$ ※

10 Gradient & Newton methods.

minimize $f(x) = -\sum_{i=1}^{m} \log(1-a_i^T x) - \sum_{i=1}^{n} \log(1-x_i^2)$

$x \in R^n$

$\mathrm{dom}\, f = \{x \mid a_i^T x < 1 \quad i=1,\dots,m$
$\quad |x_i| < 1 \quad i=1,\dots,n \}$

- solve it with Gradient & Newton method
- plot objective function v iteration numbers
  step length v iteration numbers
  $f - p^*$ v iteration numbers
- experiment α, β, see their effects
- test different cases.

<u>sol^n</u>

(a) Gradient method

△ recall algorithm

1 <u>given</u> a starting point $x \in \mathrm{dom}\, f$

2 <u>repeat</u>

- determine a step w/ $\Delta x := -\nabla f(x)$
- determine a step size w/ backtracking line search —①
- update $x := x + t\Delta x$

3 <u>until</u>

stopping criterion is satisfied —②

△ ① algorithm for backtracking line search

1 $\alpha \in (0, 0.5)$
$\beta \in (0, 1)$
$t := 1$

2 while $x + t\Delta x \notin \mathrm{dom}\, f$
$t := \beta t$

3 while $f(x+t\Delta x) > f(x) + \alpha t \nabla f(x)^T \Delta x$
$t := \beta t$

② stopping criterion

$\|\nabla f(x)\|_2 \le \eta$ (check after $\Delta x := -\nabla f(x)$)

△ below attach e-copy for MatLab implementation (see appendix)

Q.30 (cont'd)
sol (cont'd)

(b) Newton method

- recall algorithm

  1 given a starting point $x \in \mathrm{dom} f$

  2 repeat

  - determine a step w/ $\Delta x_{nt} := -\nabla^2 f(x)^{-1} \nabla f(x)$
  - determine a decrement w/ $\lambda^2 := \nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x)$.
  - stop if $\lambda^2/2 \le \epsilon$
  - determine a step size w/ backtracking line search
  - update $x := x + t\Delta x_{nt}$

- below attach e-copy for MatLab implementation (see Apendix)

9.32 Gauss-Newton

- implement Gauss-Newton for the following problem

$$f(x) = \frac{1}{2} \sum_{i=1}^{m} f_i(x)^2$$

where

$$f_i(x) = 0.5\, x^T A_i x + b_i^T x + 1$$

$$A_i \in S_{++}^n$$

$$b_i^T A_i^{-1} b_i \le 2$$

soln

△ $$\Delta x_{gn} = -\underbrace{\left(\sum_{i=1}^{m} \nabla f_i(x) \nabla f_i(x)^T\right)^{-1}}_{\text{① - term}} \underbrace{\left(\sum_{i=1}^{m} f_i(x) \nabla f_i(x)\right)}_{\text{② - term}}$$

- the ① term is the matrix generated by "$\nabla f_i(x)$" of each $i=1 \ldots m$
- the ② term is basically $\nabla f(x)$, which is the sum of $f_i(x) \nabla f_i(x)$ of each $i=1, \ldots m$.

△ implementation details are attached in appendix.

10.5

- $Q \succeq 0$

minimize $f(x) + (Ax-b)^T Q (Ax-b)$

subject to $Ax=b$ — ①

is the Newton step same as

- minimize $f(x)$

subject to $Ax=b$? — ②

sol^n

△ recall

Δx_{nt} for constrained quadratic problem :

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

△ problem ①

Δx_{nt} relationship :

$$\begin{bmatrix} \nabla^2 f(x) + A^T Q A & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{nt} \\ w_1 \end{bmatrix} = \begin{bmatrix} -\nabla f(x) - 2A^T Q A x + 2A^T Q b \\ 0 \end{bmatrix}$$

△ problem ②

Δx_{nt} relationship :

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{nt} \\ w_2 \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}$$

△ from ① & ②

①: $(\nabla^2 f(x) + A^T Q A) \Delta x_{nt} + A^T w_1 = -\nabla f(x) - 2A^T Q A x + 2A^T Q b$

$A^T \Delta x_{nt} = 0$

$\therefore$ ① is rewritten as

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{nt} \\ w_1 \end{bmatrix} = \begin{bmatrix} -\nabla f(x) - 2A^T Q A x + 2A^T Q b \\ 0 \end{bmatrix}$$

where ② is

$$\begin{bmatrix} \nabla f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{nt} \\ w_2 \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}$$

10.5 (cont'd)

sol^n (cont'd)

Δ where both prob ① & ② have

$A \, \Delta x_{nt} = 0$, while $w_1 \neq w_2$

$\therefore$ the two have same Δx_{nt} #

10.15(a)

Equality constrained entropy maximization

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

subject to $Ax = b$

where $\mathbf{dom} f = R^n_{++}$

$A \in R^{p \times n}$

$n = 100$

$p = 30$

and generate A randomly w/ full-rank

$\hat{x}$ randomly w/ $\hat{x}_i \in [0, 1]$

set $b = A\hat{x}$

show $f(x^{(k)}) - p^*$ versus k

observe quadratic convergence

sol^n

Δ recall Δx_{nt} for constrained problem:

$$\begin{array}{c} \\ n \\ p \end{array} \begin{array}{c} \begin{array}{cc} \overbrace{\quad\quad}^{n} & \overbrace{\quad}^{p} \end{array} \\ \begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \end{array} \begin{bmatrix} \Delta x_{nt} \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}$$

Δ implementation shown in appendix

- original problem

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

$\downarrow$ add constraint $x^T x \le R^2$

- new problem

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \le 0 \quad, i=1,\dots,m \\ & Ax = b \\ & x^T x \le R^2 \end{aligned}$$

- find $a > 0$ for which $\nabla^2(t f_0(x) + \tilde{\phi}(x)) \succeq aI$

<u>sol^n</u>

$\triangle$ for inequality constraints, we induce log-barrier and turn it into equality constraints-based optimization problem.

$\triangle$ original problem with log-barrier

$$\begin{aligned} \text{minimize} \quad & t f_0(x) + \phi(x) \\ \text{subject to} \quad & Ax = b \end{aligned}$$

$\triangle$ new problem with log-barrier

$$\begin{aligned} \text{minimize} \quad & t f_0(x) + \tilde{\phi}(x) \\ \text{subject to} \quad & Ax = b \end{aligned}$$

$\triangle$ we try to find $a > 0$ so that $\nabla^2(t f_0(x) + \tilde{\phi}(x)) \succeq aI$

$\triangle$ $\nabla^2(t f_0(x) + \tilde{\phi}(x))$

$$\Rightarrow \underbrace{\nabla^2(t f_0(x) + \phi(x))}_{①} + \underbrace{\frac{2}{R^2 - x^T x} I}_{②} + \underbrace{\frac{4}{(R^2 - x^T x)^2} x x^T}_{③}$$

$\triangle$ $t f_0(x) + \phi(x)$ is convex

$\therefore$ ① - term $\succeq 0$

$\triangle$ $\frac{2}{R^2 - x^T x} I \succeq \frac{2}{R^2} I \Rightarrow$ ②-term $\succeq \frac{2}{R^2} I$

$\triangle$ $\frac{4}{(R^2 - x^T x)^2} x x^T \succeq 0 \Rightarrow$ ③- term $\succeq 0$

$\checkmark$

11.4 (cont'd)

soln (cont'd)

△ $\therefore \nabla^2 (t f_0(x) + \tilde{\phi}(x)) \succeq \frac{2}{R^2} I$

$a = \frac{2}{R^2}$ ※

11.22

maximum volume rectangle inside a polyhedron

soln

△ from 8.16 the problem formulation is

$$\begin{aligned} \underset{u,l}{\text{maximize}} \quad & \sum_{i=1}^{n} \log(u_i - l_i) \\ \text{subject to} \quad & \sum_{j=1}^{n} (a_{ij}^{+} u_j - a_{ij}^{-} l_j) \le b_i, \quad i=1,\dots,m \end{aligned}$$

$$\Rightarrow \begin{aligned} \underset{u,l}{\text{minimize}} \quad & - \sum_{i=1}^{n} \log(u_i - l_i) \\ \text{subject to} \quad & A^{+}u - A^{-}l \preceq b \quad (A^{+}u - A^{-}l - b \preceq 0) \end{aligned}$$

⇒ w/ barrier method

$$\underset{u,l}{\text{minimize}} \quad -t \sum_{i=1}^{n} \log(u_i - l_i) - \sum_{i=1}^{m} \log([b - A^{+}u + A^{-}l]_i) \quad \text{—— ①}$$

△ where ① is now an unconstrained optimization problem.

△ solve it w/ Barrier method (implementation as shown in appendix.