Algorithms for Big Data Analysis : Homework #1

Due on March 7, 2017 at 23:59pm

Wen Zaiwen

Hu Tao

Problem 1

4.11 *Problems involving ℓ_1- and ℓ_∞-norms.* Formulate the following problems as LPs. Explain in detail the relation between the optimal solution of each problem and the solution of its equivalent LP.

(a) Minimize $\|Ax - b\|_\infty$ (ℓ_∞-norm approximation).

(b) Minimize $\|Ax - b\|_1$ (ℓ_1-norm approximation).

(c) Minimize $\|Ax - b\|_1$ subject to $\|x\|_\infty \le 1$.

(d) Minimize $\|x\|_1$ subject to $\|Ax - b\|_\infty \le 1$.

(e) Minimize $\|Ax - b\|_1 + \|x\|_\infty$.

In each problem, $A \in \mathbf{R}^{m\times n}$ and $b \in \mathbf{R}^m$ are given. (See §6.1 for more problems involving approximation and constrained approximation.)

Solution

(a).

$$
\begin{gathered}
min_{m,x,y} \text{ m} \\
\text{S.t y=Ax-b} \\
-m \le y_i \\
y_i \le m, (0 \le i \le n, \text{n is dimension of vector b})
\end{gathered}
$$

(b).

set notation y = Ax-b

the problem can be transfered to

$$
\begin{gathered}
min_{x,y} \parallel y \parallel_1 \\
\text{S.t y=Ax-b}
\end{gathered}
$$

$$\rightarrow$$

$$
\begin{gathered}
min_{x,y} \sum_{i=1}^{n} |y_i| \\
\text{S.t y=Ax-b}
\end{gathered}
$$

$$\rightarrow$$

$$
\begin{gathered}
min_{m,x} 1^T \cdot \text{ m} \\
\text{S.t Ax-b} \le \text{m} \\
\text{Ax-b} \ge \text{-m}
\end{gathered}
$$

(c).

$$
\begin{gathered}
min_{m,x,y} 1^T \cdot m \\
\text{S.t y=Ax-b} \\
-1 \le x_i \le 1 \\
-m \le y \le m \\
m \ge 0 \\
x_i \text{ is the i-th dimension of x .}
\end{gathered}
$$

(d).

$$
\begin{gathered}
min_{s,x}\ 1^T S \\
\text{S.t } -S_i \le x_i \le S_i \\
-1 \le a_k^T x_k - b_k \le 1
\end{gathered}
$$

(e).

$$
\begin{gathered}
min_{m,n,x}\ 1^T m + n \\
\text{S.t } -m \le Ax - b \le m \\
n \le a_k^T x_k - b_k \le n
\end{gathered}
$$

Problem 2

4.12 *Network flow problem.* Consider a network of n nodes, with directed links connecting each pair of nodes. The variables in the problem are the flows on each link: x_{ij} will denote the flow from node i to node j. The cost of the flow along the link from node i to node j is given by $c_{ij}x_{ij}$, where c_{ij} are given constants. The total cost across the network is

$$C = \sum_{i,j=1}^{n} c_{ij}x_{ij}.$$

Each link flow x_{ij} is also subject to a given lower bound l_{ij} (usually assumed to be nonnegative) and an upper bound u_{ij}.

The external supply at node i is given by b_i, where $b_i > 0$ means an external flow enters the network at node i, and $b_i < 0$ means that at node i, an amount $|b_i|$ flows out of the network. We assume that $\mathbf{1}^T b = 0$, *i.e.*, the total external supply equals total external demand. At each node we have conservation of flow: the total flow into node i along links and the external supply, minus the total flow out along the links, equals zero.

The problem is to minimize the total cost of flow through the network, subject to the constraints described above. Formulate this problem as an LP.

Solution

$$
\begin{gathered}
\min_X\ \sum_{i,j}^{n} C_{ij}X_{ij} \\
\text{S.t } l_{ij} \le x_{ij} \le u_{ij} \\
\text{for all i: } \sum_{j\neq i}(x_{ij} - x_{ji}) = b_i \\
1^T b = 0
\end{gathered}
$$

Problem 3

4.25 *Linear separation of two sets of ellipsoids.* Suppose we are given $K + L$ ellipsoids

$$\mathcal{E}_i = \{P_i u + q_i \mid \|u\|_2 \le 1\}, \quad i = 1, \ldots, K + L,$$

where $P_i \in \mathbf{S}^n$. We are interested in finding a hyperplane that strictly separates $\mathcal{E}_1, \ldots, \mathcal{E}_K$ from $\mathcal{E}_{K+1}, \ldots, \mathcal{E}_{K+L}$, *i.e.*, we want to compute $a \in \mathbf{R}^n$, $b \in \mathbf{R}$ such that

$$a^T x + b > 0 \text{ for } x \in \mathcal{E}_1 \cup \cdots \cup \mathcal{E}_K, \qquad a^T x + b < 0 \text{ for } x \in \mathcal{E}_{K+1} \cup \cdots \cup \mathcal{E}_{K+L},$$

or prove that no such hyperplane exists. Express this problem as an SOCP feasibility problem.

Problem 3 continued on next page...

Solution
from the problem,we can get:

$$inf_{a,b,u}\{a^T(P_iu+q_i)+b\} > 0 \text{ where i=1,..,K}$$
$$inf_{a,b,u}\{a^T(P_iu+q_i)+b\} < 0 \text{ where i=K+1,...,K+L}$$
$$inf_{a,u}\{a^T(P_iu)\} = inf_{a,u,||u||_2\leq 1} = -sup_{a,u}(-a^TP_iu)$$

according to the definition of operator norm,we can get:

$$inf_{a,u}\{a^T(P_iu)\} = -||a^TP_i||_2$$

so the problem can be fomulated into a SOCP as follows:

$$min_u \ 1$$
$$-||a^TP_i||_2 + a^Tq_i + b > 0 \text{ where i=1,..,K}$$
$$-||a^TP_i||_2 + a^Tq_i + b < 0 \text{ where i=K+1,...,K+L}$$

Problem 4

4.27 *Matrix fractional minimization via SOCP.* Express the following problem as an SOCP:

$$\begin{array}{ll} \text{minimize} & (Ax+b)^T(I+B\,\mathbf{diag}(x)B^T)^{-1}(Ax+b) \\ \text{subject to} & x \succeq 0, \end{array}$$

with $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$, $B \in \mathbf{R}^{m\times n}$. The variable is $x \in \mathbf{R}^n$.
Hint. First show that the problem is equivalent to

$$\begin{array}{ll} \text{minimize} & v^Tv + w^T\,\mathbf{diag}(x)^{-1}w \\ \text{subject to} & v + Bw = Ax+b \\ & x \succeq 0, \end{array}$$

with variables $v \in \mathbf{R}^m$, $w, x \in \mathbf{R}^n$. (If $x_i = 0$ we interpret w_i^2/x_i as zero if $w_i = 0$ and as ∞ otherwise.) Then use the results of exercise 4.26.

Solution
Assume x is fixed,and optimize ove v and w.
it becomes a quadratic problem,the optimal conditions are:

$$v = m, w = diag(x)B^Tm$$

for some m,Substituting in the equality constraint, we see that m must satisfy:

$$(I + Bdiag(x)B^T)v = Ax + b$$

finally we can get :

$$v = (I + B\cdot diag(x)\cdot B^T)^{-1}(Ax+b)$$
$$w = diag(x)\cdot B^T(I + B\cdot diag(x)\cdot B^T)^{-1}\cdot(Ax+b)$$

$$min_{v,w,x} \ t + 1^Ts$$
$$S.t \ v^Tv \leq t$$
$$x \geq 0$$
$$w_i^2 \leq s_ix_i, \text{ where i =1,...,n. n is the dimension of w.}$$

using the result of exercise 4.26,we can know:

$$v^T v \le t$$
$$\text{equals to}$$
$$|| \begin{pmatrix} 2v \\ t-1 \end{pmatrix} ||_2 \le t+1$$

so, the final problem can be represented as:

$$min_{v,w,x} \; t + 1^T s$$
$$S.t \; || \begin{pmatrix} 2v \\ t-1 \end{pmatrix} ||_2 \le t+1$$
$$x \ge 0$$
$$w_i^2 \le s_i x_i,$$

where i =1,...,n. n is the dimension of w.and $v = (I + B \cdot diag(x) \cdot B^T)^{-1}(Ax + b)$
$w = diag(x) \cdot B^T(I + B \cdot diag(x) \cdot B^T)^{-1} \cdot (Ax + b)$

Problem 5

3.11 *Monotone mappings.* A function $\psi : \mathbf{R}^n \to \mathbf{R}^n$ is called *monotone* if for all $x,\ y \in \mathbf{dom}\,\psi$,

$$(\psi(x) - \psi(y))^T(x - y) \ge 0.$$

(Note that 'monotone' as defined here is not the same as the definition given in §3.6.1. Both definitions are widely used.) Suppose $f : \mathbf{R}^n \to \mathbf{R}$ is a differentiable convex function. Show that its gradient ∇f is monotone. Is the converse true, *i.e.*, is every monotone mapping the gradient of a convex function?

Solution
because f(x) is convex function,so we have the following inequation:

$$f(x) \ge f(y) + \nabla f(y)^T(x - y)$$
$$f(y) \ge f(x) + \nabla f(x)^T(y - x)$$

do sum operation,get:

$$(\nabla f(x) - \nabla f(y))^T(x - y) \ge 0$$

inverse is not the same.as a counter example:

$$\psi(x) = \begin{pmatrix} x_1 \\ x_1 + x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 & x_2 \end{pmatrix}$$

$\psi is monotone because, we can get:$

$$(\nabla f(x) - \nabla f(y))^T(x - y) = (x - y)^T \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} (x - y) \ge 0$$

however,there does not exist a function: f such that $\psi(x) = \bigtriangledown f(x)$,because such function must satisfy:

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial \psi_1}{\partial x_2} = 0$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial \psi_2}{\partial x_1} = 1$$

which leads to contradiction

Problem 6

3.12 Suppose $f : \mathbf{R}^n \to \mathbf{R}$ is convex, $g : \mathbf{R}^n \to \mathbf{R}$ is concave, $\mathbf{dom}\, f = \mathbf{dom}\, g = \mathbf{R}^n$, and for all x, $g(x) \leq f(x)$. Show that there exists an affine function h such that for all x, $g(x) \leq h(x) \leq f(x)$. In other words, if a concave function g is an underestimator of a convex function f, then we can fit an affine function between f and g.

Solution

let me notate $(x,t) \in int\ epi\ f, t > f(x), and (y,v) \in hypo\ g, v \leq g(y)$
so $t \geq f(x) \geq g(x) \geq v$
therefore exists a hyperplane between int epi f and hypo g.
$a^T x + bt \geq t \geq a^T y + bv, where\ a \in R^n\ and\ b, c \in R$.and a,b cannot be both zero

$(x,t) \in epi\ f, and\ (y,v) \in hypo\ g$
if b=0,$a^T x \geq a^T y\ for\ any\ x, y$

a must be zero,which causes contradiction.therefore b cannot be zero

let x =y,so $bt \geq bv$
because $t \geq f(x) \geq g(x) \geq v$
so b > 0

we apply the separating hyperplane conditions to a point $(x,t) \in int\ epi\ f, and\ (y,v) = (x, g(x)) \in hypog$:

$$a^T x + bt \geq c \geq a^T x + bg(x)$$

divide by b:

$$t \geq (c - a^T x)/b \geq g(x)$$

therefore,exist a function $h(x) = (c - a^T x)/b$ lies between f and g.

Problem 7

2. Formulate the following optimization problems as semidefinite programs. The variable is $x \in \mathbf{R}^n$; $F(x)$ is defined as

$$F(x) = F_0 + x_1 F_1 + x_2 F_2 + \cdots + x_n F_n$$

with $F_i \in \mathbf{S}^m$. The domain of f in each subproblem is $\mathbf{dom}\, f = \{x \in \mathbf{R}^n \mid F(x) \succ 0\}$.

(a) Minimize $f(x) = c^T F(x)^{-1} c$ where $c \in \mathbf{R}^m$.

(b) Minimize $f(x) = \max_{i=1,\ldots,K} c_i^T F(x)^{-1} c_i$ where $c_i \in \mathbf{R}^m$, $i = 1, \ldots, K$.

(c) Minimize $f(x) = \sup_{\|c\|_2 \leq 1} c^T F(x)^{-1} c$.

(d) Minimize $f(x) = \mathbf{E}(c^T F(x)^{-1} c)$ where c is a random vector with mean $\mathbf{E}\, c = \bar{c}$ and covariance $\mathbf{E}(c - \bar{c})(c - \bar{c})^T = S$.

Solution

(a).
use the Schur complement,we can get:

$$minimize\ t$$
$$s.t \begin{pmatrix} F(x) & c \\ c^T & t \end{pmatrix} \geq 0$$

(b).
also use the Schur complement:

$$minimize\ t$$
$$s.t \begin{pmatrix} F(x) & c_i \\ c_i^T & t \end{pmatrix} \geq 0, i = 1, .., k$$

(c).
$f(x) = \lambda_{max}(F(x)^{-1}), so\ f(x) \leq t$,and only if $F(x)^{-1} \leq tI$.Using a Schur complement:

$$minimize\ t$$
$$s.t \begin{pmatrix} F(x) & I \\ I & tI \end{pmatrix} \geq 0$$

(d).TODO

Problem 8

3.12 *A matrix fractional function.*[1] Show that $X = B^T A^{-1} B$ solves the SDP

$$\begin{array}{ll} \text{minimize} & \mathbf{tr}\, X \\ \text{subject to} & \begin{bmatrix} A & B \\ B^T & X \end{bmatrix} \succeq 0, \end{array}$$

with variable $X \in \mathbf{S}^n$, where $A \in \mathbf{S}^m_{++}$ and $B \in \mathbf{R}^{m \times n}$ are given.
Conclude that $\mathbf{tr}(B^T A^{-1} B)$ is a convex function of (A, B), for A positive definite.

Solution
when A ¿ 0,from Schur complement,we can get following inequation:

$$\begin{pmatrix} A & B \\ B^T & X \end{pmatrix} \geq 0$$
$$\rightarrow x \geq B^T A^{-1} B$$
so $x = B^T A^{-1} B$ can solve the SDP.