Subject: HW 7 Convex opt. 810100511 محمدرضا آرانی
Year. 2022 Month June Date. 3

1) Minimizing a quadratic function

$$\min \quad f(x) = \frac{1}{2} x^T P x + q^T x + r \qquad ; \; P \in S^n$$

دو $P \succeq 0$ رو در نظر می‌گیریم

(a) if $P \not\succeq 0 \longrightarrow$ f is not convex $\longrightarrow$ problem is unbounded below

unconstrained Minimization

↳ if $P \not\succeq 0 \Rightarrow \exists v \;\; v^T P v < 0$, for $x = t v$, $t \to +\infty$

$$\Rightarrow f(x) = t^2 \frac{(v^T P v)}{2} + t(q^T v) + r \Rightarrow \boxed{f(x) \longrightarrow -\infty}$$

Dominant $\Rightarrow t \to +\infty$ — unbounded below

(b) suppose $P \succeq 0$ (f is convex), $Px^* = -q$ does not have a solution

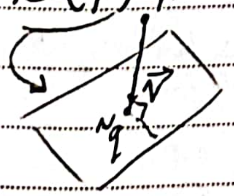
$\Rightarrow$ Problem is unbounded below یعنی که P نفی کرده که به q برسه

$\longrightarrow R(P) \neq q \longrightarrow$ یعنی q توی برد P نیست $\longrightarrow q \notin R(P)$

$\Rightarrow$ we can define a projection of q onto $R(P) \longrightarrow \tilde{q}$, $v = q - \tilde{q}$

Euclidean projection

$v \perp R(P)$



$\Rightarrow v^T P v = 0$

$$\Rightarrow t^2 \frac{(v^T P v)}{2} + t q^T v + r = t q^T v + r = t(\tilde{q} + v)^T v + r$$
$$= t \, v^T v + t \tilde{q}^T v + r$$

$\Rightarrow t \to \infty \Rightarrow f \to -\infty$

Subject:
Year. Month. Date. ( )

2) Smoothed fit to given data

$$\min f(x) = \sum_{i=1}^{n} \psi(x_i - y_i) + \lambda \sum_{i=1}^{n-1} (x_{i+1} - x_i)^2$$

$\lambda > 0$ → smoothing parameter / ψ → convex penalty function / $x \in R^n$ variable ↓ smoothed fit to vector y

(a) $$\frac{\partial}{\partial x_i} f(x) = \psi'(x_i - y_i) + \lambda \cdot \begin{cases} 2(x_1 - x_2) & i = 1 \\ 2(x_n - x_{n-1}) & i = n \\ 2(x_i - x_{i-1}) - 2(x_{i+1} - x_i) & \text{o.w.} \end{cases}$$

$$\Rightarrow \frac{\partial^2}{\partial x_i \partial x_j} f(x) = \begin{cases} j = i-1 & -2\lambda \\ j = i+1 & -2\lambda \\ i = j & \psi''(x_i - y_i) + \lambda \begin{cases} 2 & i = 1, 2 \\ 4 & \text{o.w.} \end{cases} \end{cases} \Rightarrow \nabla^2 f(x) = \begin{bmatrix} \psi''(x_1 - y_1) + 2\lambda & -2\lambda & & 0 \\ -2\lambda & \ddots & \ddots & \\ & \ddots & \ddots & -2\lambda \\ 0 & & -2\lambda & \psi''(x_n - y_n) + 2\lambda \end{bmatrix}$$

[ماتریس tridiagonal است ← به طور کلی شبیه ماتریس] to Hessian که toeplitz شبیه فرم ماتریس ✓

(b) two-dimensional : $$\min \sum_{i,j=1}^{n} \psi(x_{ij} - y_{ij}) + \lambda \left(\sum_{i=1}^{n-1} \sum_{j=1}^{n} (x_{i+1,j} - x_{ij})^2 + \sum_{i=1}^{n} \sum_{j=1}^{n-1} (x_{i,j+1} - x_{ij})^2 \right)$$

$X \in R^{n \times n}$
$Y \in R^{n \times n}$
$\lambda > 0$

$$\frac{\partial}{\partial x_{ij}} f(x) = \psi'(x_{ij} - y_{ij}) + \lambda \begin{cases} i = j = 1 & -2(x_{21} - x_{11}) - 2(x_{12} - x_{11}) \\ i = j = n & 2(x_{nn} - x_{n-1,n}) + 2(x_{nn} - x_{n,n-1}) \\ i = 1, j = n & 2(x_{1,n} - x_{1,n-1}) - 2(x_{2,n} - x_{1,n}) \\ j = 1, i = n & 2(x_{n,1} - x_{n-1,1}) - 2(x_{n,2} - x_{n,1}) \\ \text{o.w} & 2(x_{i,j} - x_{i-1,j}) + 2(x_{i,j} - x_{i,j-1}) \\ & -2[x_{i+1,j} - x_{i,j} + x_{i,j+1} - x_{i,j}] \end{cases}$$

same happens for $\frac{\partial^2 f(x)}{\partial x_{ij} \partial x_{i'j'}}$

✓ آن موقع ما یک ماتریس tridiagonal بلوکی

! با بلوک های $n \times n$ خواهیم داشت

Subject:
Year. Month. Date. ( )

3) Derive the newton Eq. for: $\min \frac{1}{2} x^T x + \log \sum_{i=1}^{m} \exp(a_i^T x + b_i)$

$A \in R^{m \times n} \to m << n$ — اول Hession objective را حساب می‌کنیم:

$\rightarrow$ rows: a_i^T

$$\frac{\partial}{\partial x_i} f(x) = x_i + \frac{\sum_{i=1}^{m} a_{mi} \cdot \exp(a_i^T x + b_i)}{\sum_i \exp(a_i^T x + b_i)}$$

$\log(u) \to \frac{u'}{u}$

$x^2 \to 2x$

$$\frac{\partial^2}{\partial x_i \partial x_j} f(x) = \begin{cases} i \neq j & \underbrace{1}_{I} + \dfrac{\sum_i a_i^2 e^{a_i^T x + b_m} \sum e^{a_i^T x + b_i} - (\sum a_i e^{a_i^T x + b_i})}{(\sum_i e^{a_i^T x + b_i})^2} \\ i = j & \to 0 + \sum a_i a_j e^{a_i^T x + b_i} \sum e^{a_i^T x + b_i} - \dfrac{\sum a_i e^{a_i^T + b_i} \cdot \sum a_i e^{a_i^T x + b_i}}{\sum_a e^{a_i^T x + b_i}} \end{cases}$$

$$\Rightarrow \nabla^2 \equiv H = \underbrace{I}_{\text{ones}(\,)} + A^T(\text{diag}(z) - zz^T)A \quad ; \quad z = \frac{e^{a_i^T x + b_i}}{\sum e^{a_i^T x + b_i}}$$

تغییر متغیر

low rank — بخش قطری‌دار به علاوه بیمارستی $\Longrightarrow$ $\text{diag}(z) - zz^T$ is singular $\Rightarrow$ نمیتوان با استفاده از Cholesky فاکتورگیری کرد

newton system: $(I + A^T(\text{diag}(z) - zz^T)A)\Delta x = g$

$\boxed{L = \text{diag}(z) - zz^T}$, $L\,\text{diag}(z)^{-1} L^T$

$$\begin{bmatrix} I & A^T L \\ L^T A & -\text{diag}(z) \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta u \end{bmatrix} = \begin{bmatrix} g \\ y \end{bmatrix}$$

$\Rightarrow \boxed{(\text{diag}(z) + L^T A A^T L)\Delta u = L^T A g}$; $\to$ m+1 متغیر $\xrightarrow[\text{محاسباتی}]{\text{هزینه}}$ $\boxed{\frac{m^3}{3} + m^2 n \text{ flops}}$

$* \ (\text{diag}(z) - zz^T)(\text{diag}(z) - zz^T)^{-1}(\text{diag}(z) - zz^T)^T = \text{diag}(z) - zz^T = L(\text{diag}(z) - zz^T)^{-1} L^T$

Subject:
Year. Month. Date. ( )

7) $X = X^T \in \mathbb{R}^{n \times n} \longrightarrow X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$; $A \in \mathbb{R}^{k \times k}$

$\det(A) \neq 0 \longrightarrow S = C - B^T A^{-1} B \rightarrow$ is Schur Complement of A in X

$\boxed{* \det(X) = \det(A) . \det(S)}$

(a) $f(u,v) = (u,v)^T X (u,v)$, $u \in \mathbb{R}^k$

$g(v) = \min_u (f) = \inf_u f((u,v))$

$f(u,v) = [u^T \ v^T] \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$

if $A \succ 0 \rightarrow g(v) = v^T S' v$

$= [\, u^T A + v^T B^T \ , \ u^T B + v^T C \,] \begin{bmatrix} u \\ v \end{bmatrix}$

نقطه اکسترمم

$= u^T A u + 2 v^T B^T u + v^T C v$

$\nabla_u f(u,v) = 2Au + 2(v^T B^T)^T = 0$

$\Rightarrow Au = -Bv \Rightarrow \boxed{u = -A^{-1}(Bv)} \Rightarrow g(v) = v^T B^T A^{-T} A A^{-1} B v - 2 v^T B^T A^{-1} B v + v^T C v$

$\rightarrow g(v) = v^T (\underline{B^T A^{-T} B} - \underline{2 B^T A^{-1} B} + C) v$; $\boxed{A^T = A \Rightarrow g(v) = v^T (C - B^T A^{-1} B) v}$

$\boxed{\Rightarrow g(v) = v^T S v}$

b) $X \succ 0$ if and only if $A \succ 0, S \succ 0$ / $X \succeq 0$ if and only if $A \succeq 0$, $B^T (I - A A^{\dagger}) = 0$, $C - B^T A^{\dagger} B \succeq 0$

$A \succ 0 \rightarrow X \succeq 0$ if and only if $S \succeq 0$ /

$\boxed{* A^{\dagger} \rightarrow \text{pseudo-inverse of } A}$

اثبات [illegible] 8

① if $X \succ 0 \Rightarrow A \succ 0, S \succ 0 \longrightarrow f(u,v) \succ 0$

$\rightarrow [u^T \ v^T] X \begin{bmatrix} u \\ v \end{bmatrix} \succ 0$

if $u = -A^{-1} B v$

$\longrightarrow \Rightarrow f(u,v) = v^T S v \succ 0 \Rightarrow \boxed{S \succ 0}$; if $v = 0 \Rightarrow f(u,v) = u^T A u \succ 0 \Rightarrow \boxed{A \succ 0}$ ✓

$\forall v$

طرف دوم

اثبات حالا if $A \succ 0, S \succ 0$ $\overset{\text{if}}{v \neq 0} \Rightarrow f(u,v) \geq \inf_u (f(u,v)) = g(v) = v^T S v \succ 0 \Rightarrow \boxed{X \succ 0}$

$\Rightarrow X \succ 0$

PAPCO

if $v = 0 \Rightarrow f(u,v) = u^T A u \rightarrow A \succ 0 \Rightarrow f(u,v) \succ 0 \Rightarrow \boxed{X \succ 0}$

Subject:
Year. Month. Date. ( )

5) Infeasible start Newton method for LP centering

$$\begin{cases} \min & c^T x - \sum_{i=1}^{n} \log(x_i) \\ \text{s.t.} & Ax = b \end{cases}$$

$*A \in R^{m \times n}$

$\begin{cases} m < n \\ c \in R^n \\ b \in R^m \end{cases}$ | A is full rank | x^*, ν^* a dual optimal point, Number of Newton steps executed

initial point $\leftarrow x^{(0)} \succeq 0$

KKT system for Newton step:
$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{nt} \\ w \end{bmatrix} = \begin{bmatrix} -g \\ 0 \end{bmatrix}$$

$g = c - (1/x_1, \dots, 1/x_n)$

$H = \operatorname{diag}\left(\frac{1}{x_1^2}, \dots, \frac{1}{x_n^2}\right)$

با روش block elimination

$$\boxed{AH^{-1}A^T w = -AH^{-1}g}$$

$\Delta x_{nt} = -H^{-1}(A^T w + g)$

optimality condition for KKT $\rightarrow A^T \nu^* + c - \left(\frac{1}{x_1^*}, \dots, \frac{1}{x_n^*}\right) = 0$ | *at convergence

$\boxed{w \text{ is the dual optimal point } \nu^*}$

6) Same happens here as happened in 5

PAPCO

Subject:
Year. Month. Date. ( )

7) estimation of a vector from one-bit measurement!

$x \in R^n$

m = sensor $\longrightarrow y_i = \text{sign}(a_i^T x + v_i - b_i) = \begin{cases} 1 & a_i^T x + v_i \geq b_i \\ -1 & a_i^T x + v_i < b_i \end{cases}$

$a_i, b_i \to$ known

$v_i \sim$ measurement error $\sim \mathcal{N}(0,1) \to \phi(v) = \left(\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{v^2}{2}}\right)$

$$P_i(x) = \text{Prob}(y_i = 1) = \frac{1}{\sqrt{2\pi}} \int_{b_i - a_i^T x}^{\infty} e^{-t^2/2} dt$$

$$1 - P_i(x) = \text{prob}(y_i = -1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{b_i - a_i^T x} e^{-\frac{t^2}{2}} dt$$

(a)
$$l(x) = \sum_{y_i = 1} \log(P_i(x)) + \sum_{y_i = -1} \log(1 - P_i(x))$$

$$P_i(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{b_i - a_i^T x}^{\infty} e^{-t^2/2} dt = \boxed{1 - F(b_i - a_i^T x)}$$

$\to$ log-concave $\leftarrow$ Gaussian dist.

Concave $\leftarrow l(x) \Leftarrow$

$\Rightarrow$ Concave maximization $\equiv$ Convex minimization ✓

(b) solve ML above / given data / newton's method with backtracking line search

$\boxed{P_i(-x) = 1 - P_i(x)}$; $A = \text{diag}(y) A$; $\longrightarrow$ b , A ضرب طرفین
$b = \text{diag}(y) b$; در مقادیر

$\Rightarrow \max \sum_{i=1}^{m} \log(\phi(Ax - b)) = f(x)$; $\longrightarrow h(\omega) = -\sum_{i=1}^{m} \log(\phi(w_i))$

$\equiv \min h(Ax - b) = f(x) \Rightarrow \nabla f(x) = A^T \nabla h(Ax - b)$; $\nabla^2 f(x) = A^T \nabla^2 h(Ax - b) A$

$\frac{\partial}{\partial w_i} h = -\frac{\phi'(w_i)}{\phi(w_i)}$

Subject:
Year. Month. Date. ( )

$$; \frac{\partial^2 h}{\partial w_i \partial w_j} = \begin{cases} \phi & i \neq j \\ -\frac{\phi''(w_i)\phi(w_i) - (\phi'(w_i))^2}{\phi^2(w_i)} & i = j \end{cases}$$

$$\frac{w_i}{\sqrt{2\pi}\cdot\left[e^{\frac{w_i^2}{2}}\right]\cdot\phi(w_i)} + \left(\frac{\frac{1}{2\pi}}{\left(e^{\frac{w_i^2}{2}}\cdot\phi(w_i)\right)^2}\right)$$

8)

PAPCO