احسان آقازاده / 810197668 / تمرین سری ۳

سوال ۳: (a)

$$f(x) = \frac{1}{2}x^T P x + q^T x + r \Rightarrow \nabla f(x) = Px + q$$

$$\Rightarrow \nabla^2 f(x) = P \quad P \not\succeq 0 \Rightarrow$$ f، convex نیست

$$P \not\succeq 0 \Rightarrow \exists v \Rightarrow v^T P v < 0$$

حال قرار می‌دهیم $x = tv$

$$\Rightarrow f(x) = \frac{1}{2}(tv)^T P(tv) + q^T(tv) + r$$

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

برای مقادیر بزرگ t، $f(x)$ به سمت $-\infty$ می‌رود پس unbounded below است.

(b)

$Px^* = -q$ این جواب ندارد $\Rightarrow -q \notin R(P) \Rightarrow q \notin R(P)$

حال q را می‌توان به شکل جمع ۲ بردار نوشت که یکی درون $R(P)$ است و دیگری بر آن عمود است.

$$\Rightarrow q = u + v \quad u \in R(P), \; v \perp R(P) \Rightarrow v^T P v = 0$$

حال قرار می‌دهیم: $x = tv$

$$\Rightarrow f(x) = \frac{t^2}{2}\underbrace{(v^T P v)}_{0} + q^T(tv) + r = t(u+v)^T v + r$$

$$= t\,\underbrace{u^T v}_{0} + t\,v^T v + r = t v^T v + r$$

← برای مقادیر بسیار منفی t، $f(x)$ به سمت $-\infty$ می‌رود پس unbounded below است.

Scanned with CamScanner

سؤال ۲ : (a)

$\nabla f^T(x) \Delta x_{nsd} =$

$\nabla f^T(x) \left(\underset{v}{\operatorname{argmin}}\{\nabla f^T(x) v \mid \|v\| \leq 1\}\right)$

$= \inf\{\nabla f^T(x) v \mid \|v\| \leq 1\}$

$= -\sup\{\nabla f^T(x) v \mid \|v\| \leq 1\}$

$= -\|\nabla f(x)\|_*$

(b)

طبق تعریف: $\Delta x_{sd} = \|\nabla f(x)\|_* \Delta x_{nsd}$

حکم: $\nabla f^T(x) \Delta x_{sd} = -\|\nabla f(x)\|_*^2$

اثبات: $\nabla f^T(x) \Delta x_{sd} = \nabla f^T(x) \|\nabla f(x)\|_* \Delta x_{nsd}$

$= \|\nabla f(x)\|_* \underbrace{\nabla f^T(x) \Delta x_{nsd}}_{(a):\ -\|\nabla f(x)\|_*} = -\|\nabla f(x)\|_*^2$ ✓

(c) میخواهیم $\underset{v}{\operatorname{argmin}}\left(\nabla f^T(x) v + \frac{1}{2}\|v\|^2\right)$ بدست آوریم. یک تغییر متغیر میدهیم.

$v = tu$ ـ شکلی که $\|u\| = 1$ حال objective ، argmin_v را بازنویسی میکنیم
t اسکالر است

$\nabla f^T(x) v + \frac{1}{2}\|v\|^2 = t \nabla f^T(x) u + \frac{t^2}{2}\cancel{\|u\|^2}$

$= t \nabla f^T(x) u + \frac{t^2}{2}$

حال برای $t \geq 0$ ، t بهینه را بدست میآوریم: (نسبت به t مشتق میگیریم و صفر قرار میدهیم)

$$\frac{\partial\left(t\nabla f^T_{(x)}u+\frac{t^2}{2}\right)}{\partial t}=0=\nabla f^T_{(x)}u+t \Rightarrow t=-\nabla f^T_{(x)}u$$

$$\Rightarrow \hat{t}=\begin{cases}-\nabla f^T_{(x)}u & \nabla f^T_{(x)}u\le 0\\ 0 & \text{o.w.}\end{cases}$$

$$\Rightarrow \hat{t}\,\nabla f^T_{(x)}u+\frac{\hat{t}^2}{2}=\frac{-(\nabla f^T(x)\,u)^2}{2}$$

داریم: $\Delta x_{nsd}=\underset{u}{\operatorname{argmin}}\{\nabla f^T_{(x)}u \mid \|u\|=1\}$ که معادل است با: $\underset{u}{\operatorname{argmin}}\{-(\nabla f^T_{(x)}u)^2/2 \mid \|u\|=1\}$

پس با کمینه کردن $-\frac{(\nabla f^T(x)\,u)^2}{2}$ برای u خواهیم داشت $\hat{u}=\Delta x_{nsd}$

$$\Rightarrow \hat{t}=-\nabla f^T(x)\,\hat{u}=-\nabla f^T(x)\,\Delta x_{nsd}=-\Delta x_{nsd}^T\nabla f^T(x)=\|\nabla f\|_*$$

$$\Rightarrow v=\hat{t}\hat{u}=\|\nabla f\|_*\,\Delta x_{nsd}=\Delta x_{sd}$$

سوال 3: (a)

$$\left[\nabla^2 f(x)\right]_{ij}=\frac{\partial^2}{\partial x_j\,\partial x_i}\Big(\sum_{t=1}^{n}\underbrace{\psi(x_t-y_t)}_{\tilde f(x_t)}+\lambda\sum_{t=1}^{n-1}\underbrace{(x_{t+1}-x_t)^2}_{g(x_t,x_{t+1})}\Big)$$

$$=\frac{\partial}{\partial x_j}\left(\frac{\partial \tilde f(x_i)}{\partial x_i}+\lambda\frac{\partial g(x_{i-1},x_i)}{\partial x_i}+\lambda\frac{\partial g(x_i,x_{i+1})}{\partial x_i}\right)$$

$$\Rightarrow \left[\nabla^2 f(x)\right]_{ij}=\begin{cases}\neq 0 & i=j\\ \neq 0 & i=j-1\\ \neq 0 & i=j+1\\ 0 & \text{o.w.}\end{cases}$$

در نتیجه H به شکل زیر خواهد شد:

$$\begin{bmatrix} a & b & & & 0\\ c & d & e & & \\ & f & g & h & \\ 0 & & & & \end{bmatrix}$$

→ به این نوع ماتریس tridiagnal گفته می‌شود

(b) این بخش هم همانند بخش قبل استدلال می‌شود و structure ماتریس H به این صورت خواهد بود که روی قطر ماتریس‌های tridiagnal وجود دارد و مابقی درایه‌ها صفر است.

سؤال $\underline{\underline{4}}$: (a)

$\hat{x}$ strictly feasible است $\Rightarrow \begin{cases} A\hat{x} = b \; * \\ \hat{x} \succeq 0 \end{cases}$

باید نشان دهیم : ① $A(\hat{x}+\Delta x) = b$ ② $\hat{x}+\Delta x \succeq 0$

① طبق معادله دوم داریم : $A\Delta x + 0.w = 0 \Rightarrow A\Delta x = 0$

$$A(\hat{x}+\Delta x) = \underbrace{A\hat{x}}_{*:\, b} + \underbrace{A\Delta x}_{0} = b + 0 = b \checkmark$$

②

$$\begin{cases} \lambda^2 = \Delta x^T \nabla^2 f(\hat{x}) \Delta x \\ \nabla^2 f(\hat{x}) = \nabla^2 \phi(\hat{x}) = \operatorname{diag}\left(\left(\frac{1}{\hat{x}_1^2}, \dots, \frac{1}{\hat{x}_n^2}\right)\right) \end{cases}$$

$$\Rightarrow \lambda^2 = \sum_{i=1}^{n} \left(\frac{\Delta x_i}{\hat{x}_i}\right)^2 \qquad \lambda \leq 1 \Rightarrow \lambda^2 \leq 1 \Rightarrow \sum_{i=1}^{n} \left(\frac{\Delta x_i}{x_i}\right)^2 \leq 1$$

$$\Rightarrow |\Delta x_i| \leq \hat{x}_i \Rightarrow \hat{x}_i + \Delta x_i \geq 0 \Rightarrow \hat{x} + \Delta x \succeq 0 \checkmark$$

(b)

باید ثابت کنیم که $A^T y \preceq c$ به ازای $y = -\frac{1}{t} w$ برقرار است

$$A^T y \preceq c \Rightarrow c - A^T y \succeq 0 \Rightarrow \underbrace{c + \frac{1}{t} A^T w \succeq 0}$$ حکم

اثبات : طبق معادله اول داریم :

$$\nabla^2 \phi(\hat{x}) \Delta x + A^T w = -tc - \nabla \phi(\hat{x})$$

$$\Rightarrow tc + A^T w = -(\nabla\phi(\hat{x}) + \nabla^2\phi(\hat{x})\Delta x) \qquad [\nabla\phi(x)]_i = \frac{-1}{x_i}$$

$$\Rightarrow c + \frac{1}{t}A^T w = \frac{-1}{t}(\nabla\phi(\hat{x}) + \nabla^2\phi(\hat{x})\Delta x)$$

$$[\nabla^2\phi(x)]_{ij} = \begin{cases} \frac{1}{x_i^2} & i=j \\ 0 & o.w. \end{cases}$$

$$\left[c + \frac{1}{t}A^T w\right]_i = \frac{-1}{t}\left(\frac{-1}{\hat{x}_i} + \frac{\Delta x_i}{\hat{x}_i^2}\right)$$

$$= \frac{1}{t\hat{x}_i}\left(1 - \frac{\Delta x_i}{\hat{x}_i}\right)$$

$$1 - \frac{\Delta x_i}{\hat{x}_i} \geq 0 \quad \Leftarrow \quad |\Delta x_i| \leq x_i$$ از قسمت قبل داشتیم که

$$\Rightarrow \left[c + \frac{1}{t}\hat{A}^T w\right]_i \geq 0 \Rightarrow c + \frac{1}{t}A^T w \succeq 0$$

$\Leftarrow$ حکم اثبات شد.

(c)

$$c^T\hat{x} - p^* \leq \underbrace{c^T\hat{x} - b^T y}$$

فاصله objective های مربوط به primal و dual

$$c^T\hat{x} - b^T y = \hat{x}^T(c - A^T y) = \frac{1}{t}\sum_{i=1}^{n}\left(1 - \frac{\Delta x_i}{\hat{x}_i}\right) = \frac{1}{t}\left(n - \underbrace{\sum_{i=1}^{n}\frac{\Delta x_i}{\hat{x}_i}}\right)$$

حد بالای این می‌شود $\sqrt{n}\,\lambda(\hat{x})$

$$v = \left(\frac{\Delta x_1}{\hat{x}_1}, \frac{\Delta x_2}{\hat{x}_2}, \ldots, \frac{\Delta x_n}{\hat{x}_n}\right) \qquad |1^T v| = \sum_{i=1}^{n}\frac{\Delta x_i}{\hat{x}_i}$$

طبق کوشی شوارتز: $|1^T v| \leq \sqrt{n}\,\underbrace{\|v\|_2}$ $\Rightarrow \left|\sum_{i=1}^{n}\frac{\Delta x_i}{\hat{x}_i}\right| \leq \sqrt{n}\,\lambda$

طبق قسمت (a) این λ می‌شود

$$\Rightarrow c^T\hat{x} - p^* \leq \frac{n + \lambda(\hat{x})\sqrt{n}}{t}$$

سؤال ۴ (a) : 5

با توجه به این که نقطه ابتدائی feasible است ، KKT سیستم به صورت زیر خواهد بود :

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

$$\nabla f(x) = c - \left(\frac{1}{x_1}, \frac{1}{x_2}, \dots, \frac{1}{x_n}\right)$$

$$\nabla^2 f(x) = \operatorname{diag}(x)^{-2}$$

$$\nabla^2 f(x)\, \Delta x_{nt} + A^T \omega = -\nabla f(x)$$

$$\Rightarrow \Delta x_{nt} = -(\nabla^2 f(x))^{-1} (A^T \omega + \nabla f(x)) \quad *$$

با توجه به معادله دوم $(A \Delta x_{nt} = 0)$ ، Δx_{nt} را از معادله اول eliminate میکنیم :

$$A(\nabla^2 f(x))^{-1} (\nabla^2 f(x)\, \Delta x_{nt}) + A(\nabla^2 f(x))^{-1} A^T \omega = -A(\nabla^2 f(x))^{-1} \nabla f(x)$$

$$\Rightarrow \underbrace{A \Delta x_{nt}}_{0} + A(\nabla^2 f(x))^{-1} A^T \omega = -A(\nabla^2 f(x))^{-1} \nabla f(x)$$

$$\Rightarrow A(\nabla^2 f(x))^{-1} A^T \omega = -A(\nabla^2 f(x))^{-1} \nabla f(x)$$

این باید حل شود و ω به دست میآید

طبق شرط KKT :

$$\nabla L(x^*, \nu^*) = 0$$

$$\Rightarrow c - \left(\frac{1}{x_1^*}, \frac{1}{x_2^*}, \dots, \frac{1}{x_n^*}\right) + A^T \nu^* = 0$$

$$(*) : \Delta x_{nt} \to 0 \Rightarrow A^T \omega + \nabla f(x) \to 0$$

$$\Rightarrow \nu^* = \omega$$

Scanned with CamScanner