حل تمرین شماره ۳ بهینه سازی

سمیه عالی

۱- الف)

$$\min \sum_{i=1}^{m} \phi(r_i) \qquad r = Ax - b$$

$$\mathcal{L}(x, r, \nu) = \sum_{i=1}^{m} \phi(r_i) + \nu^T (Ax - b - r)$$

$$g(\nu) = \min_{x, r} (\mathcal{L}(x, r, \nu)) = \min_{r} \Big(\underbrace{\min_{x} \Big(\sum_{i=1}^{m} \phi(r_i) - \nu^T b - \nu^T r + A^T \nu x \Big)}_{g_1(r, \nu)} \Big)$$

$$= \min_{r} \left(g_1(r, \nu) = \begin{cases} \sum_{i=1}^{m} \phi(r_i) - \nu^T r - \nu^T b & A^T \nu = 0 \\ -\infty & A^T \nu \neq 0 \end{cases} \right)$$

$$\xRightarrow{A^T \nu = 0} g(\nu) = \min \Big(\sum_{i=1}^{m} \phi(r_i) - \sum_{i=1}^{m} \nu_i r_i - b^T \nu \Big) = \min \sum_{i=1}^{m} \phi(r_i) - \nu_i r_i - b^T \nu$$

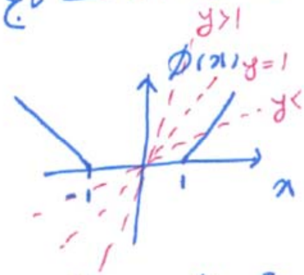
$$= \sum_{i=1}^{m} \underbrace{-\sup(\nu_i r_i - \phi(r_i))}_{\text{تابع conj}} - b^T \nu = \sum_{i=1}^{m} -\phi^*(\nu_i) - b^T \nu$$

dual problem:
$$\max \sum -\phi^*(\nu_i) - b^T \nu \qquad (I)$$
$$\text{s.t.} \quad A^T \nu = 0$$

ب) صرفاً کافی است تابع conj را برای $\phi(x)$ داده شده به دست آوریم و در (I) جایگذاری می‌کنیم:

$$\phi(x) = \begin{cases} 0 & |x| \le 1 \\ |x| - 1 & |x| > 1 \end{cases}$$

$$\phi^*(y) = \sup(yx - \phi(x))$$



با توجه به شکل مشخص است که $\phi^*(y)$ که فاصله max خط yx با $\phi(x)$ است برای $0 \le y \le 1$ ← برابر y است و برای $y > 1$ برابر ∞ می‌شود. مشابهاً این تعبیر برای y های منفی هم تکرار می‌شود لذا داریم

$$\phi^*(y) = \begin{cases} |y| & |y| \le 1 \\ \infty & |y| > 1 \end{cases}$$

لذا داریم

$$\max \sum |\nu_i| - b^T \nu \qquad \equiv \qquad \max \|\nu\|_1 - b^T \nu$$
$$\text{s.t.} \quad A^T \nu = 0, \ |\nu_i| \le 1 \qquad\qquad \text{s.t.} \quad A^T \nu = 0, \ \|\nu\|_\infty \le 1$$

۷.۳. مشاهدات (u_i, y_i) $i=1,\dots,N$ مشاهدات را طوری مرتب می‌کنیم که N_1 تای اول $y=1$ باشند و بقیه $y=-1$
لذا LF به صورت زیر است.

$$LF = \prod_{i=1}^{N_1} p(a^T u_i + b + v \leq 0) \prod_{i=N_1+1}^{N} p(a^T u_i + b + v > 0)$$

$$p(\underbrace{a^T u_i + b + v}_{z} \leq 0) = G\left(\frac{0 - (a^T u_i + b)}{1}\right) \qquad p(a^T u_i + b + v > 0) = 1 - G(0 - (a^T u_i + b))$$

$$z \sim N(a^T u_i + b, 1)$$

$$G(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-x^2/2} dx \quad \rightarrow \quad LLF = \sum_{i=1}^{N_1} \log(G(-(a^T u_i + b))) + \sum_{i=N_1+1}^{N} \log(1 - G(-a^T u_i - b))$$

۷.۵.

$$LF = P_{k_2 k_1} P_{k_3 k_2} \cdots P_{k_N k_{N-1}} = \prod_{i,j=1}^{n} P_{ij}^{m_{ij}} \qquad \begin{bmatrix} P_{11} & \cdots & P_{1j} & \cdots \\ & & P_{2j} & \\ & & \vdots & \\ & & P_{nj} & \end{bmatrix}$$

جمع ستونهای ماتریس برابر یک

$$\max \prod_{i,j=1}^{n} P_{ij}^{m_{ij}} \quad \Rightarrow \quad \max \sum_{i,j=1}^{n} m_{ij} \log(P_{ij})$$

$$\sum_{i=1}^{n} P_{ij} = 1 \quad j = 1, 2, \dots, N \qquad \sum_{i=1}^{n} P_{ij} = 1 \quad j = 1, 2, \dots, N$$

$$\Rightarrow \max \sum_{i=1}^{n} m_{ij} \log(P_{ij}) \xrightarrow[\text{شرط}]{KKT} \sum_{i=1}^{n} P_{ij} = 1$$

$$\sum_{i=1}^{n} P_{1j} = 1$$

$$\nabla \mathcal{L}(P, v) = 0$$

$$\mathcal{L} = \sum_{i=1}^{n} m_{ij} \log(P_{ij}) + v\left(\sum_{i=1}^{n} P_{ij} - 1\right) \Rightarrow \begin{bmatrix} \frac{m_{1j}}{P_{1j}} \\ \vdots \\ \frac{m_{nj}}{P_{nj}} \end{bmatrix} + \begin{bmatrix} v \\ v \\ \vdots \\ v \end{bmatrix} = 0$$

$$\Rightarrow P_{ij} = \frac{-m_{ij}}{v} = \frac{m_{ij}}{\sum m_{ij}}$$

$$\sum P_{ij} = 1 = \sum -\frac{m_{ij}}{v} = 1 \Rightarrow v = -\sum_i m_{ij}$$

مقاله

$\Rightarrow g(\ldots$

الف)

(۱/۵)

$$p=\begin{bmatrix} \frac{1}{6} & \frac{1}{12} \\ \frac{1}{9} & \frac{1}{6} \\ 0 & \frac{1}{6} \\ \frac{1}{6} & 0 \\ \frac{1}{9} & 0 \\ \frac{1}{4} & \frac{1}{3} \\ 0 & \frac{1}{4} \\ \frac{1}{12} & 0 \\ 0 & \frac{1}{12} \\ \frac{1}{9} & \frac{1}{12} \end{bmatrix}$$
$$\underbrace{}_{p}\quad\underbrace{}_{q}$$

$$D=TP=\begin{bmatrix} t_{11} & t_{12} & \cdots & t_{110} \\ t_{21} & t_{22} & \cdots & t_{210}\end{bmatrix}\times[p\ q]$$

$$\min\ P_{Fa},\ P_{FN}=(TP)_2,\ (Tq)_1 \quad (۰/۵)$$

$$s.t \quad t_{1k}+t_{2k}=1 \qquad k=1,\ldots,10 \quad (۱)$$

$$t_{ik}\geq 0 \qquad i=1,2,\ k=1,\ldots,10 \quad (۰/۵)$$

$$(TP)_2=\frac{1}{6}t_{21}+\frac{1}{9}t_{22}+0\times t_{23}+\cdots+\frac{1}{9}t_{2,10} \quad (۲/۵)$$

$$(Tq)_1=\frac{1}{12}t_{11}+\frac{1}{6}t_{12}+\cdots+\frac{1}{12}t_{1,10} \quad (۲/۵)$$

ب) (۳/۵)

حل ساده $\Longrightarrow (t_{1k},t_{2k})=\begin{cases}(1,0) & p_k\geq q_k\\ (0,1) & p_k<q_k\end{cases}$

$$\min\ (TP)_2+(Tq)_1$$

$$s.T:\quad t_{1k}+t_{2k}=1,\ t_{ik}\geq 0,\ i=1,2,\ k=1,\ldots,10$$

$$T=\begin{bmatrix}1&0&0&1&1&1&0&1&0&1\\0&1&1&0&0&0&1&0&1&0\end{bmatrix}\Longrightarrow$$

نوع ML است

(۱/۵)

$$\min \|x\|_2^2 \qquad \mathcal{L}(x,\lambda) = x^T x + \lambda \times -(a^T x + b)$$ ٤-

$$-(a^T x + b) \leq 0 \qquad \nabla \mathcal{L}(x,\lambda) = 2x - \lambda a = 0 \implies x = \frac{\lambda a}{2} \quad ①$$

complementary slackness

$$\lambda \times (a^T x + b) = 0$$

$$\lambda = 0 \overset{①}{\implies} x = 0 ,\ \ -b \leq 0 \implies b \geq 0$$

$$\lambda \neq 0 \implies a^T x + b \overset{①}{=} \lambda \frac{a^T a}{2} + b = 0 \implies \lambda = \frac{-2b}{\|a\|_2^2} \overset{①}{\implies} x = \frac{-ba}{\|a\|_2^2}$$

$$\lambda \geq 0 \implies -b < 0$$

٥-

$$\mathcal{L}(x, y, \nu) = 0.5 \|y - y_0\|_2^2 + \sum_{i=1}^m \|x_i\|_2 + \sum_{i=1}^m \nu_i^T (x_i - C_i y - d_i)$$

$$g(\nu) = \min_y \Big(\underbrace{\min_x \sum_{i=1}^m \|x_i\|_2 + \nu_i^T x_i}_{\sum_{i=1}^m -\sup(-\|x_i\|_2 + \nu_i^T x_i)} + 0.5\|y - y_0\|_2^2 + \sum_{i=1}^m -\nu_i^T C_i y + \sum_{i=1}^m -\nu_i^T d_i \Big)$$

$$= 0 \quad \|\nu_i\|_2 \leq 1 \quad i = 1, \dots, m$$

$$= -\infty \quad \text{otherwise}$$

$$g(\nu) = \min_y \Big(0.5 \|y - y_0\|_2^2 + \sum_{i=1}^m -\nu_i^T C_i y + \sum_{i=1}^m -\nu_i^T d_i \Big)$$

$$i = 1, \dots, m \quad \|\nu_i\|_2 \leq 1$$

$$\nabla_y = 0 \qquad \cancel{2}(y - y_0) - \sum C_i^T \nu_i = 0 \implies y = y_0 + \sum C_i^T \nu_i$$

$$g(\nu) = \begin{cases} 0.5 \|\sum C_i^T \nu_i\|_2^2 & \\ -\infty & \text{otherwise} \end{cases}$$

٦-

$$\min f(x) \quad Ax = b \implies \mathcal{L}(x,\nu) = f(x) + \nu^T (Ax - b)$$

$$\nabla \mathcal{L}(x,\nu) = \nabla f(x) + A^T \nu \quad ①$$

$$\arg\min_x f(x) + a\|Ax - b\|_2^2 = x^* \implies \nabla f(x^*) + 2aA^T(Ax^* - b) = 0 \quad ②$$

به ازای ①,② $\nu^* = 2a(Ax^* - b)$ ، x^* ، $\mathcal{L}(x,\nu)$ را min می‌کند پس ν^* feasible است

$$g(\nu^*) = f(x^*) + \nu^{*T}(Ax^* - b) = f(x^*) + 2a\|Ax^* - b\|_2^2$$

۷- شرایط KKT را برای مسئله اصلی می‌نویسیم.

$$\mathcal{L}(x,\lambda_1,\lambda_2) = -c^Tx - \frac{1}{2}x^TQx + \lambda_1^T(Ax-b) - \lambda_2^Tx$$ ۳

$$\nabla\mathcal{L}(x,\lambda_1,\lambda_2) = 0 = -c - Qx + A^T\lambda_1 - \lambda_2 \Rightarrow$$ ۳

شرط ۴ KKT

با مقایسه با آنچه می‌خواهیم $\lambda_1 = u$ و $\lambda_2 = y$.

شرط ۱: $x \geq 0$, $Ax-b \leq 0$ ① ۱

شرط ۲: $u \geq 0$ و $y \geq 0$ ۱

شرط ۳: ③ $u_i(a_i^Tx - b_i) = 0$ ۱ $\Rightarrow v_i = -(a_i^Tx - b_i)$ ۲ ②

④ $y_i x_i = 0$ ۱

$\Rightarrow$ داریم $x \geq 0$ ، $y \geq 0$ ، $u \geq 0$ و $v \geq 0$ طبق ① و ②

پس رابطه ③ و ④ را می‌توان به فرم $x^Ty + u^Tv = 0$ نوشت چون همه متغیرها مثبتند ۳

از طرفی از تعریف v هم رابطه دومی که می‌خواهیم به دست می‌آید.

۸-

$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} \Delta x \\ w \end{bmatrix} = \begin{bmatrix} -g \\ 0 \end{bmatrix} \xrightarrow[\text{جدید داریم}]{\text{برای مسئله}} \begin{bmatrix} H + A^TQA & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} \Delta x \\ w \end{bmatrix} = \begin{bmatrix} -g - 2A^TQAx + 2A^TQb \\ 0 \end{bmatrix}$$ ۵

حال چون داریم $A\Delta x = 0$ پس رابطه بالا به فرم زیر ساده می‌شود

$A^TQA\Delta x = 0$

$$\Rightarrow \begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} \Delta x \\ w \end{bmatrix} = \begin{bmatrix} \cdot \\ 0 \end{bmatrix} \xrightarrow{\hat{w} = w + 2QAx - 2Qb} \begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} \Delta x \\ \hat{w} \end{bmatrix} = \begin{bmatrix} -g \\ 0 \end{bmatrix}$$ ۴

پس Δxها با گام نیوتن همان گام نیوتن برای مسئله $Q=0$ شد.

با ترم quad اضافه شده، جهت‌ها ثابت‌اند و از آنجایی که با ساده فرض کرده‌ایم فرض‌ها $Ax=b$ یعنی xهایی که برای مسئله feasible شده همگی در primal quad اضافه شده مستقل لذا جواب یکی می‌شوند. ۴