احسان آقازاده/ 810197668/ تمرین سری پنجم

سؤال ۱:

a) این مسئله را با روش تخمین بیشینه درست‌نمایی حل می‌کنیم. باید درست‌نمایی (likelihood) را بیشینه کنیم. likelihood خروجی مسابقات به شرط a که مقدار توانایی تیم‌ها است به صورت زیر است:

$$P(y|a) = \prod_{i=1,\dots,m} \Phi\left(\frac{1}{6} y^{(i)} \left(a_{j(i)} - a_{k(i)}\right)\right)$$

در این رابطه Φ، cdf توزیع نرمال استاندارد است و به این دلیل که نمونه‌ها یا همان بازی‌ها i.i.d هستند، می‌توان احتمال آن‌ها را در هم ضرب نمود. حال از این عبارت لگاریتم می‌گیریم تا log-likelihood بدست بیاید.

$$l(a) = \log(P(y|a)) = \sum_{i=1}^{m} \log\left(\Phi\left(\frac{1}{6} y^{(i)} \left(a_{j(i)} - a_{k(i)}\right)\right)\right)$$

$$= \sum_{i=1}^{m} \log\left(\Phi\left(\left(\frac{1}{6}\right)(Aa)_i\right)\right)$$

$l(a)$ به این دلیل که جمع تعدادی لگاریتم است (هر کدام از آن‌ها concave است)، تابعی concave است. حال مسئله به شکل زیر خواهد بود:

$$\text{maximize} \quad l(a)$$
$$\text{subject to} \quad 0 \preceq a \preceq 1$$

که این مسئله با توجه به اینکه مسئله maximization است و $l(a)$ تابعی concave

است، یک مسئله convex optimization است. دقت کنید شرط $0 \le a_i \le 0$ از آنجا می‌آید که توانایی هر شیم در بازه $[0,1]$ است.

سؤال ج:

چون $0<l<u$ و برای هر مقدار t، $f'(t) \in [l,u]$

$\to$ تابع f وارون‌پذیر است *

$$y_i = f(a_i^T x + b_i + v_i), \quad i=1,\dots,m$$

$$\overset{*}{\to} \quad v_i = f^{-1}(y_i) - a_i^T x - b_i \quad i=1,\dots,m$$

P_v را cdf مربوط به توزیع متغیر تصادفی v در نظر می‌گیریم. حال likelihood را به شکل زیر می‌نویسیم:

$$P(x,f \mid y) = \prod_{i=1}^{m} P_v(f^{-1}(y_i) - a_i^T x - b_i)$$

حال log-likelihood به شکل زیر می‌شود:

$$l(x,f) = \sum_{i=1}^{m} \log\left(P_v(f^{-1}(y_i) - a_i^T x - b_i)\right)$$

$$= \sum_{i=1}^{m} \log\left(P_v(\tilde{x}_i - a_i^T x - b_i)\right)$$

که $\tilde{x}_i = f^{-1}(y_i)$،

چون $l(x,f)$ مجموع تعدادی لگاریتم است (که هرکدام concave هستند) پس این تابع نسبت به x و $\tilde{x}$، concave است.

Scanned with CamScanner

$$f'(t) \in [l, u] \Rightarrow (f^{-1})'(t) \in [1/u, 1/l]$$

$$\left(\frac{1}{u}\right)|y_i - y_j| \leq |\tilde{x}_i - \tilde{x}_j| \leq (1/l)\,|y_i - y_j|$$

برای هر i و j ای باید برقرار باشه

که این در واقع شرط (constraint) مسئله خواهد بود

در نتیجه چون می‌خواهیم likelihood را maximize کنیم با توجه به concave بودن تابع و وجود نامساوی‌های خطی (linear inequality) یک مسئله convex optimization خواهیم داشت که به شکل زیر است:

$$\text{maximize} \quad l(x, \tilde{x}) = \sum_{i=1}^{m} \log\left(p_v(\tilde{x}_i - a_i^T x - b_i)\right)$$

$$\text{subject to} \quad \left(\frac{1}{u}\right)|y_i - y_j| \leq |\tilde{x}_i - \tilde{x}_j| \leq (1/l)\,|y_i - y_j|$$

$$i, j = 1, \ldots, m$$

سوال 3: (a)

$$y_i = \phi(a_i^T x + v_i), \quad i = 1, \ldots, m$$

$$\Rightarrow v_i = \underbrace{\phi^{-1}(y_i)}_{z_i} - a_i^T x, \quad i = 1, \ldots, m$$

با توجه به اینکه $\alpha \leq \phi'(u) \leq \beta$ و $0 < \alpha < \beta$ برقرار است $\Leftarrow$ تابع ϕ وارون‌پذیر است

Scanned with CamScanner

حال از روش تخمین بیشینه درست‌نمایی استفاده می‌کنیم. v از توزیع نرمال استاندارد است

$$P(v) = (2\pi\sigma^2)^{-1/2} e^{-v^2/(2\sigma^2)}$$

با توجه به این که نویزها i.i.d هستند

$$P(v|y) = (2\pi\sigma^2)^{-m/2} \prod_{i=1}^{m} e^{-(\phi^{-1}(y_i) - a_i^T x)^2/2\sigma^2}$$

$$= (2\pi\sigma^2)^{-\frac{m}{2}} \prod_{i=1}^{m} e^{-(z_i - a_i^T x)^2/2\sigma^2}$$

log-likelihood

$$l(x,z) = \frac{-m}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{m}(z_i - a_i^T x)^2$$

این تابع یک term ثابت (constant) دارد و یک term که برابر منفی جمع تعدادی quadratic (البته composition آن با affine که همچنان quadratic و convex باقی خواهد ماند) است پس $l(x,z)$ یک تابع concave است و maximize کردن آن نیز یک مسئله convex optimization است. دقت شود که یک constraint هم داریم که به صورت زیر است:

$$\frac{1}{\beta} \le (\phi^{-1})'(u) \le \frac{1}{\alpha}$$

ترتیب yها به شکل nondecreasing است

$$\frac{y_{i+1} - y_i}{\beta} \le z_{i+1} - z_i \le \frac{y_{i+1} - y_i}{\alpha} \quad i=1,\dots,m$$

Scanned with CamScanner

مسئله را به شکل خلاصه می‌توان به صورت زیر نوشت :

$$\text{maximize} \quad -\sum_{i=1}^{m} (z_i - a_i^T x)^2$$

$$\text{subject to} \quad \frac{y_{i+1}-y_i}{\beta} \le z_{i+1}-z_i \le \frac{y_{i+1}-y_i}{\alpha} \quad i=1,\dots,m$$

سوال 5 :

ellipsoid $\mathcal{E}$ شامل ellipsoid $\mathcal{E}_i$ خواهد بود اگر شرط زیر برقرار باشد :

$$\sup_{x \in \mathcal{E}_i} (x-x_0)^T A^{-1} (x-x_0) \le 1$$

$$x^T A_i x + 2 b_i^T x + c_i \le 0 \Rightarrow x^T A^{-1} x - 2 x_0^T A^{-1} x + x_0^T A^{-1} x_0 - 1 \le 0$$

$$(x \in \mathcal{E}_i)$$

با توجه به ~~تضمینست~~ appendix B، این عبارت درست خواهد بود اگر و تنها اگر $\lambda_i \ge 0$ موجود باشد که :

$$\lambda_i \begin{bmatrix} A_i & b_i \\ b_i^T & c_i \end{bmatrix} \succeq \begin{bmatrix} A^{-1} & -A^{-1}x_0 \\ -(A^{-1}x_0)^T & x_0^T A^{-1} x_0 - 1 \end{bmatrix}$$

$$\begin{bmatrix} \lambda_i A_i & \lambda_i b_i \\ \lambda_i b_i^T & 1+\lambda_i c_i \end{bmatrix} - \begin{bmatrix} I \\ -x_0^T \end{bmatrix} A^{-1} \begin{bmatrix} I & -x_0 \end{bmatrix} \succeq 0$$

$$\begin{bmatrix} A & I & -x_0^T \\ I & \lambda_i A_i & \lambda_i b_i \\ -x_0 & \lambda_i b_i^T & 1+\lambda_i c_i \end{bmatrix} \succeq 0 \longrightarrow$$

این شرط به ازای هر $\mathcal{E}_i$ خواهد بود

Scanned with CamScanner

می‌خواهیم $\log(\det A^{-1})$ را minimize کنیم. می‌توان w.l.o.g فرض کرد $A \in S_{++}$ است پس $A^{-1} \in S_{++}$ است پس $\log(\det A^{-1})$، convex است و یک مسئله‌ی convex optimization از نوع SDP خواهیم داشت که به شکل زیر است:

$$\begin{aligned} &\text{minimize} \quad \log(\det A^{-1}) \\ &\text{subject to} \quad \begin{bmatrix} A & I & -x_0 \\ I & \lambda_i A_i & \lambda_i b_i \\ -x_0^T & \lambda_i b_i^T & 1+\lambda_i c_i \end{bmatrix} \succeq 0, \quad i=1,\dots,k \\ &\qquad \lambda_i \geq 0 \quad i=1,\dots,k \end{aligned}$$

سؤال 5: ابتدا اثبات می‌کنیم که حداقل یک نقطه باشد که projection x_0 بر روی C باشد:

یک نقطه مانند $\tilde{x}$ درون C فرض می‌کنیم ($\tilde{x} \in C$). حال اگر مقدار $\|x - x_0\|$ را بر روی $C \cap \{x \mid \|x-x_0\| \leq \|\tilde{x}-x_0\|\}$، minimize کنیم، نقطه‌ی مورد نظر را به دست خواهیم آورد. با توجه به اینکه $\|x-x_0\|$ convex و feasible set هم intersection، C که convex است و $\{x \mid \|x-x_0\| \leq \|\tilde{x}-x_0\|\}$ که sublevel set یک تابع convex است و convex می‌شود پس کل feasible set، convex می‌شود و مسئله convex optimization و در نتیجه می‌شود حداقل یک نقطه را داریم.

حال اثبات می‌کنیم که این نقطه منحصر به فرد است:

از برهان خلف استفاده می‌کنیم:

فرض می‌کنیم x_1^* و x_2^* نقاط projection x_0 بر روی C هستند ($x_1^*, x_2^* \in C$)

پس $\Leftarrow$ $\|x_1^*-x_0\|=\|x_2^*-x_0\|=d$

نقطه‌ی $x=\frac{x_1^*+x_2^*}{2}$ را در نظر می‌گیریم $\Leftarrow$ $x\in C$

$$\|x-x_0\| = \|\tfrac{1}{2}(x_1^*-x_0)+\tfrac{1}{2}(x_2^*-x_0)\|$$

$$\frac{1}{2}\underbrace{\|x_1^*-x_0\|}_{d}+\frac{1}{2}\underbrace{\|x_2^*-x_0\|}_{d}=d$$ $<$ طبق قواعد نرم‌ها

$\Leftarrow$ نقطه‌ای پیدا شد درون C که فاصله‌اش با x_0 از فاصله‌ی x_1^* و x_2^* با x_0

کمتر است پس نقاط x_1^* و x_2^* نقاط projection نیستند که در تناقض با فرض است

$\Leftarrow$ اثبات شد که این نقطه projection بر روی C منحصر به فرد است

سوال ۷:

① $t^*>0 \Rightarrow$ دو مجموعه نقاط linearly جداشدنی هستند

اثبات:

$$t^*>0 \Rightarrow a^{*T}x_i \geq t^*+b^* > b^* > b^*-t^* \geq a^{*T}y_i$$

در نتیجه با a^* و b^* یک hyperplane جداکننده تعریف می‌شود

$\Rightarrow$ نقاط به صورت linearly جداشدنی هستند

② دو مجموعه نقاط linearly جداشدنی هستند $\Rightarrow t^*>0$

Scanned with CamScanner

اثبات :

دو مجموعه نقاط linearly جداپذیر هستند $\Longrightarrow$ a و b ای وجود دارد که بتوان با آن hyperplane جداکننده تعریف کرد

t مثبتی وجود دارد که با constraint ها برقرار باشند

با توجه به اینکه مسئله maximization است پس $t^* > 0$

~~constraint~~

$\|a\|_2 \le 1$ constraint ای tight است به این دلیل که constraint های دیگر homogeneous هستند .

(ب) فرض می‌کنیم a و b و t یک عضو از feasible set مجموعه 8.23 باشند .

شرایط زیر برای a و b و t برقرار خواهد بود.

$$\left.\begin{array}{ll} a^T x_i - b \ge t & i = 1, \dots, N \\ a^T y_i - b \le -t & i = 1, \dots, M \\ \|a\|_2 \le 1 & \end{array}\right\} *$$

اگر دو مجموعه جداشدنی باشند ، $t > 0$ [redacted]

$$\xrightarrow{\;*\;} \quad \frac{a^T}{t} x_i - \frac{b}{t} \ge 1 \quad i = 1, \dots, N$$

$$\frac{a^T}{t} x_i - \frac{b}{t} \le -1 \quad i = 1, \dots, M$$

$\Leftarrow$ $\tilde{a} = \frac{a}{t}$ و $\tilde{b} = \frac{b}{t}$ داخل feasible set مسئله QP است و objective نیز

Scanned with CamScanner

برابر $\|a\|_2 / t = \|\tilde{a}\|_2$ خواهد بود.

حال فرض می‌کنیم $\tilde{a}$ و $\tilde{b}$ یک عضو از feasible set مسئله QP باشد

$$\Rightarrow \quad \tilde{a}^T x_i - \tilde{b} \geq 1 \quad i=1,\dots,N$$

$$\tilde{a}^T y_i - \tilde{b} \leq -1 \quad i=1,\dots,M$$

$$\frac{\tilde{a}^T}{\|\tilde{a}\|_2} x_i - \frac{\tilde{b}}{\|\tilde{a}\|_2} \geq \|\tilde{a}\|_2 \quad i=1,\dots,N$$

$$\frac{\tilde{a}^T}{\|\tilde{a}\|_2} y_i - \frac{\tilde{b}}{\|\tilde{a}\|_2} \leq -\|\tilde{a}\|_2 \quad i=1,\dots,M$$

پس با تعریف $a = \frac{\tilde{a}}{\|\tilde{a}\|_2}$ و $b = \frac{\tilde{b}}{\|\tilde{b}\|_2}$ و $t = \frac{1}{\|\tilde{a}\|_2}$ و چون $\|a\|_2 = \frac{\|\tilde{a}\|_2}{\|\tilde{a}\|_2} = 1$ داریم

a و b و t یک عضو از feasible set مسئله 8.23 خواهد بود و objective نیز برابر $t = \frac{1}{\|\tilde{a}\|_2}$ است

$\Leftarrow$ این مسئله QP با مسئله 8.23، equivalent است.