810100441 hw5 حمید رضا کاشانی

Q-1 مشله a) L norm است که باید شروط بهینگی را بنویسیم برای این کار می‌توانیم $\frac{3}{2}$ تابع را minimize کنیم:

مسئله روبه رو را حل می‌کنیم:

$$\text{minimize } f(x) = \sum_{i=1}^{m} |a_i^T x - b_i|^{3/2}$$

تابع مشتق پذیر است بنابراین شرط بهینگی این است که گرادیان صفر شود:

$$\nabla f = \sum_{i=1}^{m} \left(\frac{3}{2}\right) \operatorname{sign}(a_i^T x - b_i)\, \underbrace{|a_i^T x - b_i|}_{r_i}^{1/2} a_i$$

پس داریم:

$$\sum_{i=1}^{m} \left(\frac{3}{2}\right) \operatorname{sign}(r_i) |r_i|^{1/2} a_i = 0$$

b) مسئله بالا را به فرم SDP می‌خواهد؛ که می‌توانیم یک مسئله معادل برای مسئله خود بنویسیم به صورت زیر:

$t, s \in R^m$

$$\text{minimize } 1^T t$$
$$\text{s.t } s^{3/2} \preceq t, \quad -s_i \preceq a_i^T x - b_i \preceq s_i \quad i = 1, \dots, m$$

که باید بتوانیم $s^{3/2} \leq t$ را به صورت LMI دربیاوریم:

$$s_i^{3/2} \leq t_i \rightarrow \frac{s_i^2}{\sqrt{s_i}} \leq t_i \rightarrow s_i^2 \leq t_i \sqrt{s_i}$$

که $\sqrt{s_i}$ خطی نیست

$$\xrightarrow{LMI} \begin{bmatrix} \sqrt{s_i} & s_i \\ s_i & t \end{bmatrix} \geq 0$$

متغیر جدید y را تعریف می‌کنیم به صورت زیر:

$y \in R^m$ و $0 \leq y \leq \sqrt{s}$

$$\rightarrow \underbrace{\begin{bmatrix} y_i & s_i \\ s_i & t_i \end{bmatrix} \succeq 0}_{\substack{y_i \geq 0 \\ y_i t_i \geq s_i^2}}, \quad \underbrace{\begin{bmatrix} s_i & y_i \\ y_i & 1 \end{bmatrix} \succeq 0}_{s_i \geq y_i^2}$$

$$\sqrt{s_i} \geq |y_i| \rightarrow \sqrt{s_i} \geq y_i$$

$$\rightarrow \sqrt{s_i} \geq y_i, \quad y_i t_i \geq s_i^2 \Rightarrow \boxed{\sqrt{s_i}\, t_i \geq s_i^2}$$

مسئله SDP ⇐

$$\text{minimize } 1^T t$$
$$\text{s.t } -s_i \leq a_i^T x - b_i \leq s_i \quad i = 1, \dots, m$$
$$\begin{bmatrix} y_i & s_i \\ s_i & t_i \end{bmatrix} \succeq 0 \quad \begin{bmatrix} s_i & y_i \\ y_i & 1 \end{bmatrix} \succeq 0 \quad i = 1, \dots, m$$

c) کد python زده شده است.

2 با استفاده از hint داده شده، ما از توابع لاگرانژین پایه $f_0, \dots, f_K$ استفاده می‌کنیم:

$$f_0(x) = \left(\frac{a_1 - x}{a_1 - a_1}\right)_+ , \quad f_i(x) = \left(\min\left(\frac{x - a_{i-1}}{a_i - a_{i-1}}, \frac{a_{i+1} - x}{a_i - a_{i+1}}\right)\right)_+ \quad i = 1, \dots, K-1$$

$$f_K(x) = \left(\frac{x - a_{K-1}}{a_K - a_{K-1}}\right)_+$$

تابع f را به صورت زیر parametrized کرده:

$$f(x) = \sum_{i=0}^{K} z_i f_i(x) , \quad z_i = f(\alpha_i) , \quad z = (z_0, \dots, z_K)$$

The least-squares fitting criterion:

$$J = \sum_{i=1}^{m} (f(x_i) - y_i)^2 = \|Fz - y\|_2^2$$

که $F \in R^{m \times (K+1)}$

که F ماتریسی است به طوری که: $F_{ij} = f_j(x_i)$, $i = 1, \dots, m$, $j = 0, \dots, K$

و باید قید محدب بودن f را وارد کنیم ← slope های هر قسمت صعودی است پس داریم:

$$\frac{z_{i+1} - z_i}{a_{i+1} - a_i} \geq \frac{z_i - z_{i-1}}{a_i - a_{i-1}} \quad i = 1, \dots, K-1 \implies$$ نامساوی خطی

بهترین مسئله کانوکس piecewise linear:

$$\text{minimize} \quad \|Fz - y\|_2^2$$

$$\text{s.t.} \quad \frac{z_{i+1} - z_i}{a_{i+1} - a_i} \geq \frac{z_i - z_{i-1}}{a_i - a_{i-1}} \quad i = 1, \dots, K-1$$

که با قیود نابرابری اضافه شده است.

3- سؤال SVM چند لیبل است یعنی خروجی چندین حالت دارد و داریم:

$$L(A,b) = \sum_{i=1}^{m} \left(1 + \max_{k \neq y_i} f_k(x_i) - f_{y_i}(x_i)\right)$$

و مسئله:

$$\text{minimize} \quad L(A,b) + \mu \|A\|_F^2$$

$$\text{s.t.} \quad 1^T b = 0 \quad , \mu > 0$$

a) باید ثابت کنیم مسئله کانوکس است:

- $\|A\|_F^2$ یک تابع quadratic است، convex است، تابع f_K خطی است پس $\max f_K$ بیشینه یک سری توابع خطی است که آن هم کانوکس می‌شود تابع ۲ هم affine است پس جمع یک سری توابع convex که کانوکس می‌شود پس تابع بهینه‌سازی convex است و شرط هم affine است پس مسئله ما convex است.

b) کد ضمیمه شده است.

4 - a) می‌خواهیم مسئله فرموله کنیم با استفاده از maxlikelihood ماتریس A ماتریسی است با ستون‌ها به تعداد بازی و سطر‌ها به تعداد تیم‌ها یک سطر بازی برده مثبت یک سطر بازنده منفی یک عضو دیگر آن بازی به تیم دیگر ندارد: (i,j)

$$A_{ij} = \begin{cases} y^{(i)} & (i,j) \\ -y^{(i)} & j=k \;(i) \\ 0 & ow \end{cases}$$

خروجی تیم‌ها به شرط داشتن a ها برابر است با:

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

$\longrightarrow$ Φ تابع تجمعی توزیع نرمال است

حال تابع log-likelihood برابر است:

$$l(a) = \log(p(y|a)) = \sum_i \log \Phi\left(\left(\frac{1}{\sigma}\right)(Aa)_i\right)$$

$\longrightarrow$ این یک تابع concave است

$$\max \; l(a)$$
$$\text{s.t} \;\; 0 \le a \le 1$$

max likelihood را تخمین می‌زند $\hat{a}$ مسئله روبه‌رو است:

که مسئله بیشینه کردن یک تابع concave با شروط نامساوی خطی است به تعداد $2n$ نامساوی خطی پس مسئله convex است و قابل حل

b) فایل python

c) فایل python به همراه جواب در انتهای آن

---

5 - قرار است مسئله را فرمولایز کنیم و به صورتی که مسئله کانوکس باشد تا loss را از آن محاسبه کنیم. برای این می‌گیریم که $P \in R_+^{n \times n}$ ماتریس توزیع احتمال‌ها است به طوری که $P_{ij} = \text{Prob}(R_1 = r_i, R_2 = r_j)$ شرایطی که marginals ها یکی داده شده باشد:

$$P^T 1 = p^{(2)} \quad , \quad P1 = p^{(1)}$$

شرط correlation هم می‌تواند به این صورت بیان گردد:

$$(r - \mu_1 1)^T P (r - \mu_2 1) = \rho \sigma_1 \sigma_2$$

احتمال ضرر $R_1 + R_2 \le 0$ از رابطه روبه‌رو به دست می‌آید:

$$\text{Prob}(R_1 + R_2 \le 0) = \sum_{r_i + r_j \le 0} P_{ij}$$

بنابراین مسئله LP به صورت زیر است:

$$\text{maximize} \; \sum_{r_i + r_j \le 0} P_{ij}$$

$$\text{s.t} \;\; P_{ij} \ge 0 \; ; \; i,j = 1,\dots,n \;\; , \;\; P1 = p^{(1)} \;\; , \;\; P^T 1 = p^{(2)} \;\; , \;\; (r - \mu_1 1)^T P (r - \mu_2 1) = \rho \sigma_1 \sigma_2$$

حال حل به صورت python ضمیمه شده است و در آن دقیقا حدود 0.192 درآمده است که چهار برابر R1 و R2 گفته شده توزیع است و توزیع حاصل رکورد رسم شده است.

a-6) می‌خواهیم مسئله را تعریف کنیم، می‌بینیم برای x ول ثابت شده بهترین انتخاب برای y_j برای A_{ij} هایی که داریم این است که y_j را با A_{ij}/x_i برابر بگیریم زیرا در آن صورت برای این درایه‌ها ($R(A_{ij}, x_i, y_j)$ صفر می‌شود. پس می‌توانیم فرض کنیم که متغیر بهینه‌سازی فقط x ها ول است و داریم:

$$\text{minimize} \sum_{(i,j)\in\Omega} R(A_{ij}, x_i, y_j)$$

$$\text{s.t.} \quad x \succeq 0,\ y \succeq 0,\ 1^T x = 1 \qquad R(u,v) = \max\left(\frac{u}{v}, \frac{v}{u}\right) - 1$$

می‌بینیم که تابع R کانوکس نیست پس باید سعی کنیم با تغییر متغیرهایی ببینیم که می‌شود آن را کانوکس کرد یا نه پس ما $u = \log x$ و $v = \log y$ را در نظر گرفتیم $x = e^u$ و $y = e^v$ حال داریم:

$$R(A_{ij}, x_i, y_j) = \max\left(\frac{A_{ij}}{e^{u_i}e^{v_j}}, \frac{e^{u_i}e^{v_j}}{A_{ij}}\right) - 1 = \max\left(e^{-u_i - v_j + \log A_{ij}}, e^{u_i + v_j - \log A_{ij}}\right) - 1$$

$$= e^{|u_i + v_j - \log A_{ij}|} - 1$$

حال چون ۱- یک عدد ثابت است پس می‌توان $\Rightarrow$ از تابع بهینه‌سازی آن را حذف کرد $\rightarrow$ پس تابع بهینه‌سازی می‌شود $\Leftarrow$

$$\boxed{\sum_{(i,j)\in\Omega} e^{|u_i + v_j - \log A_{ij}|}}$$

زیرا $\log A_{ij}$ که یک عدد ثابت می‌شود و $|u_i + v_j - \log A_{ij}|$ قدر مطلق affine است و این هم کانوکس می‌شود حال exp گرفتن از کانوکس چون صعودی است تابع کانوکس می‌دهد و جمع چند کانوکس تابع کانوکس است.

حال این کار را کردیم که بتوانیم متغیر را affine کنیم زیرا با ضرب α در x ها و ضرب $\frac{1}{\alpha}$ در y ها مسئله تغییر نمی‌کند پس داریم:

$$u = \log x \Rightarrow \log \alpha x = \log \alpha + \log x = \log x + \beta \Rightarrow u \to u + \beta$$

$$\log y/\alpha = -\log \alpha + \log y = \log y - \beta \Rightarrow v \to v - \beta$$

حال شرط ما که $1^T x = 1$ بود با تغییر متغیرمان از affine بودن خارج شده بود به شکل زیر می‌شود:

$$1^T x = 1^T e^u = 1 \to \underset{\text{نیست}}{\text{affine}} \Rightarrow e^{u_1} + \dots + e^{u_n} = e^{u_1+\beta} + \dots + e^{u_n+\beta}$$

$$\Rightarrow e^{\beta}(e^{u_1} + \dots + e^{u_n}) \Rightarrow$$ e^{β} را می‌توان طوری انتخاب کرد که

حال می‌توان برای جمع u_i ها برابر صفر همین β را پیدا کرد که شرط را برابر ۱ کند پس مسئله تبدیل شد به:

$$\text{minimize} \sum_{(i,j)\in\Omega} e^{|u_i + v_j - \log A_{ij}|}$$

$$\text{s.t.} \quad 1^T u = 0$$

این مسئله کانوکس است و به ما u و v های بهینه را می‌دهد که از آن x و y ها را می‌شود حساب کرد اما مشکل این است که به جای جمع x_i ها برابر یک، ضرب آن‌ها برابر یک است که می‌توان ترفند زیر را برای $\hat{x}$ و $\hat{y}$ به دست آمده از حل مسئله زد تا نرمالایز شوند

$$\boxed{x^* = \hat{x}/(1^T\hat{x}) \qquad y^* = (1^T\hat{x})\,\hat{y}}$$