810100441 hw5 حمید رضا کاشانی

Q-1 مسئله ما L norm ۱.۵ است که باید شروط بهینگی را بنویسیم برای این کار می‌توانیم ۳/۲ تابع را minimize کنیم:

مسئله رو به رو را حل می‌کنیم:

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

تابع مشتق‌پذیر است بنابراین شرط بهینگی این است که گرادیان صفر شود:

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

پس داریم:

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

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

$t, s \in R^m$

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

که باید سطر $s^{3/2} \le t$ را به صورت LMI دربیاوریم: $s_i^{3/2} \le t_i \to \frac{s_i^2}{\sqrt{s_i}} \le t_i \to s_i^2 \le t_i \sqrt{s_i}$ که $\sqrt{s_i}$ خطی نیست

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

متغیر جدید y را تعریف می‌کنیم به صورتی که: $y \in R^m$ و $0 \le y \le \sqrt{s}$

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

$s_i \ge y_i^2$

$\sqrt{s_i} \ge |y_i| \to \sqrt{s_i} \ge y_i$

$$\hookrightarrow \sqrt{s_i} \ge y_i,\ y_i t_i \ge s_i^2 \Longrightarrow \boxed{\sqrt{s_i}\, t_i \ge s_i^2}$$

مسئله SDP $\Longleftarrow$

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

c) کد در python زده شده است.

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

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

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

تابع f را به صورت زیر parametrized کرده:

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

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 ازای $j = 0, \ldots, K$ و $i = 1, \ldots, m$: $F_{ij} = f_j(x_i)$

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

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

بهترین مسئله کانوکس piecewise linear:

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

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

که با قیود نامساوی 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 تابع f_K خطی است پس $\max f_K$ بیشینه یک سری توابع خطی است که آن هم کانوکس می‌شود تابع 1 هم affine است پس جمع یک سری توابع convex کانوکس می‌شود ~~[illegible]~~ پس تابع بهینه‌سازی conex است و شرط هم affine است پس مسئله ما convex است.

b) در صفحه بعد است.

4- a) می‌خواهیم مسئله فرموله کنیم با استفاده از maxlikelihood : ماتریس A ماتریسی است با ستون‌ها به تعداد بازی و سطر ها برابر با تعداد تیم‌ها. یک یعنی در بازی برده منفی یک یعنی باخته صفر یعنی آن بازی به تیم دیگر ندارد: (i,j)

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

خروجی تیم‌ها به شرط داشتن a ها برابر است با:

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

$\rightarrow$ Φ تابع تجمعی توزیع نرمال است

حال تابع log-likelihood برابر است:

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

$\rightarrow$ concave این یک تابع است

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

$$\max \; l(a)$$
$$s.t \quad 0 \le a \le 1$$

که مسئله بیشینه کردن یک تابع concave با شروط نامساوی خطی است که 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 فرموله کنیم و به صورتی که مسئله کانوکس باشد تا loss را از آن محاسبه کنیم. برای آن $P \in R_+^{n \times n}$ می‌گیریم که P 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:

$$P1 = p^{(1)} \quad , \quad P^T1 = p^{(2)}$$

شرط correlation ها 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:

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

$$s.t \quad P_{ij} \ge 0 ; \; i,j = 1, \dots, n \; , \; P1 = p^{(1)} \; , \; P^T1 = p^{(2)} \; , \; (r - \mu_1 1)^T P (r - \mu_2 1) = \rho \sigma_1 \sigma_2$$

حال حل به صورت python ضمیمه شده است و در آن دیده می‌شود حدود 0.192 درآمده است که جهار برابر R1 و R2 گوسی توزیع است و توزیع حاصل رسم شده است.