HW05 - liner algebra

Saeed Rafati

98106542

سوال اول

با استفاده از الگوریتم MP و OMP مسئله P_0 را برای x حل نمایید. همچنین مقدار $spark(A)$ و $\mu(A)$ را محاسبه کنید.

$$y = Ax$$

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \end{bmatrix} ، y = \begin{bmatrix} 0 \\ 2 \\ 3 \\ 5 \end{bmatrix}$$

① برای بدست آوردن $spark(A)$ می‌توانیم متغیرهای آزاد دستگاه را مشخص کنیم.

سپس هر بار دستگاه را به ازای اینکه یک متغیر آزاد 1 باشد و بقیه 0 باشند حل می‌کنیم. آن جوابی

که بیشترین $x_i = 0$ را داشت برابر $spark(A)$ می‌باشد.

← اگر $rref(A)$ را در بیاوریم، سپس دستورالعمل بالا را اجرا کنیم به $spark(A)=4$ می‌رسیم.

$$rref(A) = \left[\begin{array}{c|c} \underbrace{\begin{matrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{matrix}}_{A'} & \underbrace{\begin{matrix} 3 \\ 5 \\ -3 \\ 0 \end{matrix}}_{y'} \end{array}\right]$$

$$x_{p_1} = [2, 4, -2, 0, 1, 0]^T$$

$$x_{p_2} = [2, 4, -3, 1, 0, 1]^T$$

جواب‌های ویژه ←

← همانطور که مشاهده می‌شود به ازای x_{p_1}، $\boxed{spark(A)=4 \text{ می‌شود}}$

برای محاسبه‌ی $\mu(A)$ ابتدا A را نرمالیزه می‌کنیم $(\tilde{A})$. سپس ماتریس $(\tilde{A})^T(\tilde{A})$ را حساب می‌کنیم و بزرگترین درایه‌ای که روی قطر اصلی نباشد، برابر $\mu(A)$ می‌باشد.

$$\tilde{A} = \begin{bmatrix} \frac{1}{\sqrt2} & 0 & \frac{1}{\sqrt2} & 0 & 0 & \frac{1}{\sqrt2} \\ 0 & \frac{1}{\sqrt2} & \frac{1}{\sqrt2} & \frac{1}{\sqrt2} & 0 & 0 \\ \frac{1}{\sqrt2} & 0 & 0 & \frac{1}{\sqrt2} & \frac{1}{\sqrt2} & 0 \\ 0 & \frac{1}{\sqrt2} & 0 & 0 & \frac{1}{\sqrt2} & \frac{1}{\sqrt2} \end{bmatrix}$$

$$\rightarrow \max(\tilde{A}^T A) = \boxed{\mu(A) = 0.5 \Rightarrow 1 + \frac{1}{\mu(A)} = 3}$$

① MP algorithm

(K=1) $e_j = \dfrac{\|r^{K-1}\|^2 - (a_j^T r^{K-1})^2}{a_j^T a_j} = \begin{cases} 33.5 \\ 13.5 \\ \vdots \\ 6 \checkmark \\ 25.5 \end{cases}$

$\Rightarrow S^1 = \{5\},\ z_5^* = 4$

$r^1 = [0, 2, -1, 1]$ و $\|r^1\|_2 = 6$

← گام‌های بعدی را نیز به همین روند طی می‌کنیم داریم:

(K=2), $e_j = \begin{cases} 5.5 \\ 1.5 \checkmark \\ 4 \\ 5.5 \\ 6 \\ 5.5 \end{cases} \Rightarrow S^2 = \{2, 5\} \qquad z_2^* = 1.5$

$r = [0, 0.5, -1, -0.5]^T,\ \|r^2\|_2 = 1.5$

(K=3), $e_j = \begin{cases} 1.5 \\ 1.375 \\ 1.375 \\ 0.375 \checkmark \\ 1.375 \end{cases} \rightarrow S^3 = \{2, 5, 5\},\ r^3 = [0, 0.5, -0.25, 0.25]^T$ و $\|r^3\|_3 = 0.375$

← باید با همین روند پیش رفت تا بالاخره در مرحله K ام، به شرط $\|r^K\|_2 < \epsilon$ برسیم.

② OMP algorithm

(K=1) $\rightarrow e(j) = \begin{cases} 33.5 \\ 13.5 \\ 36 \\ 25.5 \\ 6 \checkmark \xrightarrow{} \text{index} = 5 \\ 25.5 \end{cases}$

$\Rightarrow S^1 = \{5\}, \quad A_S = [a_5]$

$(A_S^T A_S)^{-1}(A_S^T b) = 4$

$r^1 = b - 4\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$

(K=2): $\rightarrow e(j) = \begin{cases} 5.5 \\ 1.5 \checkmark \rightarrow \text{index} = 2 \\ 4 \\ 5.5 \\ 6 \\ 5.5 \end{cases}$

$\Rightarrow S^2 = \{2, 5\} \Rightarrow A_S^2 = [a_2\ a_5]$

$\rightarrow (A_S^{2T} A_S)^{-1}(A_S^{2T} b) = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

$\rightarrow r^2 = b - \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix} = 0 \Rightarrow$ الگوریتم پایان می‌پذیرد.

سوال سوم

در الگوریتم OMP در گام شماره k ، مجموعه S به صورت $S^k = S^{k-1} \cup \{j_0\}$ به روز رسانی میشود که j_0 ستونی از ماتریس واژه نامه $A \in (R)^{m\times n}$ است که $\epsilon(j)$ تعریف شده به صورت زیر را حداقل میکند:

$$||\mathbf{r}^{k-1}||_2^2 - \frac{(\mathbf{a}_j^T \mathbf{r}^{k-1})^2}{||\mathbf{a}_j||_2^2}$$

نشان دهید برای بردار دلخواه $\mathbf{b}$ و تکرار k ،مجموعه S^k در دوحالت زیر یکسان است:

- ماتریس واژه نامه $\mathbf{A} = \begin{bmatrix} \mathbf{a_1} & \mathbf{a_2} & \cdots & \mathbf{a_n} \end{bmatrix}$
- ماتریس واژه نامه $\widetilde{\mathbf{A}} = \mathbf{AD}$ که در آن $\mathbf{D} \in \mathbb{R}^{n\times n}$ یک ماتریس گطری معکوس پذیر است.

سوال چهارم

الف. بردار دلخواه $\mathbf{b} \in \mathbb{R}^m$ را در نظر بگیرید. می‌خواهیم بردار $\mathbf{x} \in \mathbb{R}^n$ در همسایگی b را به گونه‌ای به دست آوریم که نسخه هموارشده آن باشد. این مسئله را به صورت حداقل مربعات چندهدفه زیر مدلسازی نمایید ($\lambda > 0$):

$$\min_{\mathbf{x}} ||\mathbf{A}\mathbf{x} - \mathbf{b}||^2 + \lambda ||\mathbf{D}\mathbf{x} - \mathbf{d}||^2$$

که در مسئله فوق عبارت اول نزدیکی دو بردار $\mathbf{x}$ و $\mathbf{b}$ و عبارت دوم، همواری بردار $\mathbf{x}$ را برآورده می‌نماید (ماتریس‌های $\mathbf{A}$، $\mathbf{D}$ و $\mathbf{d}$ را به دست آورید که $\mathbf{D} \in \mathbb{R}^{(n-1)\times n}$ است).

ب. برای حالت خاص $n = 2$، $\hat{\mathbf{x}}(\lambda)$ (پاسخ مسئله بهینه‌سازی فوق) را به دست آورید.

ج. بردار $\lim_{\lambda\to\infty} \hat{\mathbf{x}}(\lambda)$ را به دست آورید و توجیه کنید.

(a)

چون می‌خواهیم فاصله‌ی x تا b را مینیمم کنیم بهترین انتخاب برای A، ماتریس همانی می‌باشد $\Leftarrow$ $A = I_{n\times n}$

همچنین برای هموارسازی، اختلاف درایه‌های پشت سرهم را کمترین کنیم یعنی

$$D = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & & & & \vdots \\ 0 & 0 & 1 & -1 & \cdots & & & \vdots \\ \vdots & & & & & & \ddots & \\ 0 & \cdots & & & & & 1 & -1 \end{bmatrix}_{(n-1)(n)}$$

همچنین برای هموارسازی باید $Dx-d$ مینیمم شود که $\Leftarrow$ Dx مینیمم شود $\Leftarrow$ یک انتخاب خوب برای $d = [0]_{n-1}$

(b) $n=2 \Rightarrow D = [1, -1]$, $d = 0$, $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

می‌دانیم $\hat{x}(\lambda) = \underbrace{(\tilde{A}^T \tilde{A})^{-1}} \tilde{A}^T \tilde{b}$, where $\tilde{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \sqrt{\lambda} & -\sqrt{\lambda} \end{bmatrix}$, $\tilde{b} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix}$

$$(\tilde{A}^T \tilde{A})^{-1} = \left(\begin{bmatrix} 1 & 0 & \sqrt{\lambda} \\ 0 & 1 & -\sqrt{\lambda} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \sqrt{\lambda} & -\sqrt{\lambda} \end{bmatrix} \right)^{-1} = \frac{1}{1+2\lambda} \begin{bmatrix} 1+\lambda & \lambda \\ \lambda & 1+\lambda \end{bmatrix} \quad \text{(I)}$$

$$\xrightarrow{\text{(I)}} \hat{x}(\lambda) = \frac{1}{1+2\lambda} \begin{bmatrix} 1+\lambda & \lambda \\ \lambda & 1+\lambda \end{bmatrix} \begin{bmatrix} 1 & 0 & \sqrt{\lambda} \\ 0 & 1 & -\sqrt{\lambda} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$\Rightarrow \boxed{\hat{x}(\lambda) = \frac{1}{1+2\lambda} \left(b_1 \begin{bmatrix} 1+\lambda \\ \lambda \end{bmatrix} + b_2 \begin{bmatrix} \lambda \\ 1+\lambda \end{bmatrix} \right) = \frac{1}{1+2\lambda} \begin{bmatrix} (1+\lambda) b_1 + \lambda b_2 \\ \lambda b_1 + (1+\lambda) b_2 \end{bmatrix}}$$

(c)

$$\lim_{\lambda\to\infty} \hat{x}(\lambda) = \begin{bmatrix} \frac{b_1+b_2}{2} \\ \frac{b_1+b_2}{2} \end{bmatrix}$$

سوال پنجم

در مسئله حداقل مربعات، تابع هدف ما (جهت کمینه کردن)، به صورت زیر است:

$$||Ax-b||^2 = \sum_{i=1}^{m}(\tilde{a}_i^T - b_i)^2$$

که در آن $\tilde{a}_i^T$ سطر iام ماتریس A است و هدف، یافتن بردار $\hat{x} \in \mathbb{R}^n$ است. در یک نسخه دیگر از مسئله حداقل مربعات، تابع هدف ما به صورت زیر است:

$$\sum_{i=1}^{m} w_i(\tilde{a}_i^T - b_i)^2$$

که وزن‌های w_i مثبت هستند و به ما داده شده‌اند.

الف. نشان دهید با انتخاب ماتریس قطری D به صورت مناسب، می‌توان این مسئله را به مسئله حداقل مربعات معمولی تبدیل کرد به طوری که تابع هدف به صورت $||D(Ax-b)||^2$ خواهد شد و عملا تابع هدف در مسئله جدید خواهد بود:

$$||Bx-d||^2$$

به طوری که $B = DA$ و $d = Db$.

ب. نشان دهید اگر ستون‌های ماتریس A مستقل خطی باشند، ستون‌های ماتریس B نیز مستقل خطی هستند.

چ. جواب مسئله حداقل مربعات جدید را بر حسب A و b و $W := \mathbf{diag}(w)$ به دست آورید.

(a)

فرض کنید برای A داریم -

$$A = \begin{bmatrix} - & a_1^T & - \\ - & a_2^T & - \\ & \vdots & \\ - & a_m^T & - \end{bmatrix}$$

اگر D را تعریف کنیم:

$$D = \begin{bmatrix} \sqrt{w_1} & & & 0 \\ & \sqrt{w_2} & & \\ & & \ddots & \\ 0 & & & \sqrt{w_m} \end{bmatrix}$$

به آن نگاه داریم:

$$B = DA = \begin{bmatrix} - & \sqrt{w_1}\, a_1^T & - \\ - & \sqrt{w_2}\, a_2^T & - \\ & \vdots & \\ - & \sqrt{w_m}\, a_m^T & - \end{bmatrix}, \quad d = Db = \begin{bmatrix} \sqrt{w_1}\, b_1 \\ \vdots \\ \sqrt{w_m}\, b_m \end{bmatrix}$$

$$\xrightarrow{\text{حال داریم}} ||Bx-d||^2 = \sum_{i=1}^{m}\left(\sqrt{w_i}\, a_i^T x_i - \sqrt{w_i}\, b_i\right)^2$$

$$\Rightarrow \sum_{i=1}^{m}(\sqrt{w_i})^2\left[a_i^T x_i - b_i\right]^2 \Rightarrow \boxed{\sum_{i=1}^{m} w_i (a_i^T x_i - b_i)^2}$$

(b)

از برهان خلف استفاده می‌کنیم. فرض کنیم که ستون های B مستقل خطی نباشند

$Bx = 0 \rightarrow$ یک جواب غیر بدیهی غیر صفر خواهد داشت.

$$\Rightarrow DAx = 0 \xrightarrow{\times D^{-1}} Ax = 0 \rightarrow x$$

به تناقض رسیدیم. درحالیکه فرض کرده بودیم ستون های A مستقل خطی می‌باشند.

ج

← می‌دانیم جواب معادلات مربعات $||Ax-b||$ از رابطه‌ی $x^*=(A^TA)^{-1}Ab$ بدست می‌آید .

← همینطور اثبات کردیم که جواب $||Bx-d||$ هم از جواب $\sum_{i=1}^{m} w_i(x_i-b)^2$ می‌باشد

$$\Rightarrow x^*=(B^TB)^{-1}Bd \xrightarrow[\substack{D=\sqrt{diag(W)}\ \text{II} \\ d=Db\ \text{III}}]{B=DA\ \text{I}} x^*=(A^T\underbrace{D^TD}_{Diag(W)}A)^{-1}DADb$$

$$\Rightarrow \boxed{x^* = (A^T diag(W) A)^{-1} A\, diag(W)\, b}$$