به نام خداوند بخشنده مهربان

شبکه های اجتماعی و اقتصادی

دانشکده مهندسی کامپیوتر
مریم رمضانی

بهار ۱۴۰۴

آمار و احتمال، جبرخطی

کوییز اول

۱۲ اسفند ۱۴۰۳

۱. هر کوییز 0.5 نمره از بارم کلی درس دارد.

۲. می توانید بدون حل یک سوال، از نتیجه اش در حل سوالات دیگر استفاده نمایید.

پرسش ۱ (۲۰ نمره)

(آ) اثبات کنید عبارت $E[(X-c)^2]$ زمانی مینیمم میشود که $c=\mu$ باشد.

(ب) اثبات کنید عبارت $E[|X-c|]$ زمانی مینیمم می شود که c برابر با میانه X باشد.

پاسخ:

(آ) عبارت $\mathbb{E}[(X-c)^2]$ را بسط می دهیم:

$$\begin{aligned}\mathbb{E}[(X-c)^2] &= \mathbb{E}[X^2 - 2cX + c^2] \\ &= \mathbb{E}[X^2] - 2c\mathbb{E}[X] + c^2\end{aligned}$$

برای یافتن مقدار مینیمم، مشتق نسبت به c را می‌گیریم و آن را برابر با صفر قرار می‌دهیم:

$$\frac{d}{dc}\mathbb{E}[(X-c)^2] = -2\mathbb{E}[X] + 2c = 0$$

حل برای c:

$$c = \mathbb{E}[X] = \mu$$

مشتق دوم برابر است با $\frac{d^2}{dc^2}\mathbb{E}[(X-c)^2] = 2 > 0$، که تأیید می‌کند این یک مینیمم است.
بنابراین، $\mathbb{E}[(X-c)^2]$ زمانی مینیمم می‌شود که $c=\mu$ باشد.

(ب) فرض کنید $F_X(x)$ تابع توزیع تجمعی X باشد.

$$\begin{aligned}\mathbb{E}[|X-c|] &= \int_{-\infty}^{\infty} |x-c| dF_X(x) \\ &= \int_{-\infty}^{c} (c-x) dF_X(x) + \int_{c}^{\infty} (x-c) dF_X(x)\end{aligned}$$

مشتق‌گیری نسبت به c:

$$\begin{aligned}\frac{d}{dc}\mathbb{E}[|X-c|] &= \int_{-\infty}^{c} dF_X(x) - \int_{c}^{\infty} dF_X(x) \\ &= F_X(c) - (1 - F_X(c)) = 2F_X(c) - 1\end{aligned}$$

قرار دادن این عبارت برابر با صفر:

$$\begin{aligned}2F_X(c) - 1 &= 0 \\ F_X(c) &= \frac{1}{2}\end{aligned}$$

این بدان معنی است که c مقداری است که برای آن $P(X \le c) = \frac{1}{2}$، که تعریف میانه X است.
بنابراین، $\mathbb{E}[|X-c|]$ زمانی مینیمم می‌شود که c برابر با میانه X باشد.

پرسش ۲ (۲۰ نمره)

نرم فروبنیوس برای ماتریس $A \in \mathbb{R}^{m\times n}$ به صورت زیر تعریف می‌شود:

$$\|A\|_F = \sqrt{\text{tr}\{A^T A\}}$$

(آ) ثابت کنید:

$$\|A\|_F = \sqrt{\sum_{i=۱}^{m}\sum_{j=۱}^{n}|A_{ij}|^۲}.$$

(ب) اگر مقادیر تکین ماتریس A را با $\sigma_۱, \sigma_۲, \dots, \sigma_r$ نشان دهیم، ثابت کنید:

$$\|A\|_F = \sqrt{\sum_{i=۱}^{r}\sigma_i^۲}.$$

(ج) ثابت کنید:

$$\sigma_{\max}(A) \leq \|A\|_F \leq \sqrt{r}\sigma_{\max}(A).$$

پاسخ:

(آ) با تعریف شروع می‌کنیم: $\|A\|_F = \sqrt{tr\{A^T A\}}$

عنصر (i,j) ماتریس $A^T A$ عبارت است از:

$$(A^T A)_{ij} = \sum_{k=۱}^{m} A_{ki}\cdot A_{kj}$$

trace ماتریس $A^T A$ عبارت است از:

$$tr\{A^T A\} = \sum_{i=۱}^{n}(A^T A)_{ii} = \sum_{i=۱}^{n}\sum_{k=۱}^{m}A_{ki}\cdot A_{ki} = \sum_{i=۱}^{n}\sum_{k=۱}^{m}|A_{ki}|^۲$$

تغییر اندیس‌ها:

$$tr\{A^T A\} = \sum_{k=۱}^{m}\sum_{i=۱}^{n}|A_{ki}|^۲ = \sum_{i=۱}^{m}\sum_{j=۱}^{n}|A_{ij}|^۲$$

بنابراین: $\|A\|_F = \sqrt{tr\{A^T A\}} = \sqrt{\sum_{i=۱}^{m}\sum_{j=۱}^{n}|A_{ij}|^۲}$

(ب) با استفاده از تجزیه SVD داریم: $A = U\Sigma V^T$، که در آن:

- U یک ماتریس متعامد $m\times m$ است
- Σ یک ماتریس قطری $m\times n$ با مقادیر منفرد $\sigma_۱, \sigma_۲, ..., \sigma_r$ بر روی قطر است
- V یک ماتریس متعامد $n\times n$ است

جایگذاری در تعریف نرم فروبنیوس:

$$\begin{aligned}\|A\|_F &= \sqrt{tr\{A^T A\}} = \sqrt{tr\{(U\Sigma V^T)^T(U\Sigma V^T)\}}\\ &= \sqrt{tr\{V\Sigma^T U^T U\Sigma V^T\}}\end{aligned}$$

از آنجایی که U متعامد است، $U^T U = I$:

$$\|A\|_F = \sqrt{tr\{V\Sigma^T\Sigma V^T\}}$$

با استفاده از ویژگی‌های trace و اینکه V متعامد است ($V^T V = I$):

$$\|A\|_F = \sqrt{tr\{\Sigma^T\Sigma V^T V\}} = \sqrt{tr\{\Sigma^T\Sigma\}}$$

از آنجایی که $\Sigma^T\Sigma$ قطری است و $\sigma_i^۲$ روی قطر آن قرار دارد:

$$\|A\|_F = \sqrt{\sum_{i=۱}^{r}\sigma_i^۲}$$

(ج) برای نامساوی سمت چپ: از قسمت (ب) داریم: $\|A\|_F = \sqrt{\sum_{i=1}^{r} \sigma_i^2}$

از آنجایی که همه σ_i^2 نامنفی هستند و با فرض $\sigma_1 = \sigma_{max}(A)$:

$$\sum_{i=1}^{r} \sigma_i^2 \geq \sigma_1^2$$

با گرفتن جذر:

$$\|A\|_F \geq \sigma_1 = \sigma_{max}(A)$$

برای نامساوی سمت راست: از آنجایی که $\sigma_i \leq \sigma_1$ برای همه i ها:

$$\sum_{i=1}^{r} \sigma_i^2 \leq \sum_{i=1}^{r} \sigma_1^2 = r\sigma_1^2$$

با گرفتن جذر:

$$\|A\|_F \leq \sqrt{r}\sigma_1 = \sqrt{r}\sigma_{max}(A)$$

بنابراین: $\sigma_{max}(A) \leq \|A\|_F \leq \sqrt{r}\sigma_{max}(A)$

پرسش ۳ (۱۰ نمره) ماتریس های U و V در SVD Decomposition چه معنی و کاربردی دارند؟

پاسخ:

در تجزیه SVD ماتریس $A \in \mathbb{R}^{m\times n}$ به صورت $A = U\Sigma V^T$:

ماتریس U (بردارهای منفرد چپ):

- U یک ماتریس متعامد $m \times m$ است که ستون‌های آن بردارهای ویژه AA^T هستند.
- r ستون اول یک مبنای متعامد برای فضای ستونی A تشکیل می‌دهند.
- ستون‌های باقی‌مانده مبنایی برای فضای پوچ چپ A تشکیل می‌دهند.

ماتریس V (بردارهای منفرد راست):

- V یک ماتریس متعامد $n \times n$ است که ستون‌های آن بردارهای ویژه A^TA هستند.
- r ستون اول یک مبنای متعامد برای فضای ردیفی A تشکیل می‌دهند.
- ستون‌های باقی‌مانده مبنایی برای فضای پوچ A تشکیل می‌دهند.

کاربردها:

(آ) **تبدیل خطی:** U و V مبناهای متعامد را برای دامنه و هم‌دامنه فراهم می‌کنند، و Σ فاکتورهای مقیاس‌بندی را می‌دهد.

(ب) **تحلیل مولفه‌های اصلی:** ستون‌های V مولفه‌های اصلی را نشان می‌دهند (جهت‌های حداکثر واریانس).

(ج) **فشرده‌سازی داده:** برش به k مقدار منفرد برتر، بهترین تقریب رتبه-k را می‌دهد.

(د) **شبه‌وارون:** محاسبه $A^+ = V\Sigma^+U^T$ برای حل معادلات کمترین مربعات.

(ه) **تحلیل معنایی پنهان:** در پردازش زبان طبیعی، آن‌ها "مفاهیم پنهان" را در مجموعه‌های سند نشان می‌دهند.

(و) **پردازش سیگنال:** مبنای توابع متعامد را برای تجزیه سیگنال فراهم می‌کنند.

در اصل، U و V ساختار اساسی تبدیل خطی نشان داده شده توسط A را ضبط می‌کنند، که آن‌ها را برای کاهش ابعاد، تحلیل داده‌ها و حل سیستم‌های خطی ارزشمند می‌سازد.