بسمه تعالی

بازشناسی آماری الگو

دانشگاه صنعتی خواجه نصیرالدین طوسی ۱۳۰۷ — تمرین سری دوم — تاریخ تحویل: ۱۴۰۴/۷/۲۶ — آزمایشگاه بینایی ماشین و پردازش تصاویر پزشکی

۱- (اختیاری) یکی از کاربردهای قطری سازی همزمان دو ماتریس، کاهش ابعاد داده‌ها به منظور تسهیل دسته‌بندی آنها است. داده‌های متعادل (هم تعداد) کلاس فیروزه‌ای و ارغوانی را به ترتیب با میانگین $\mu_1 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ و $\mu_2 = \begin{bmatrix} -3 \\ -3 \end{bmatrix}$ و ماتریس کواریانس $\Sigma_1 = \Sigma_2 = \begin{bmatrix} 4 & 4 \\ 4 & 9 \end{bmatrix}$ در نظر بگیرید (تصویر اول).

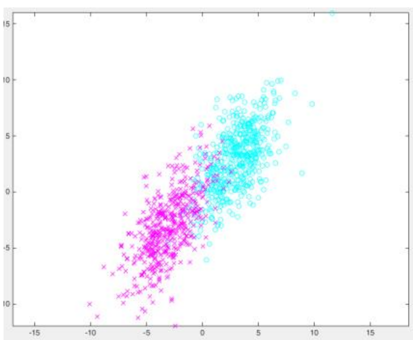
الف. ماتریس کواریانس بین کلاس‌ها را با استفاده از معادله زیر محاسبه کنید (μ معرف میانگین عمومی و μ_c معرف میانگین کلاس c است).

$$\Sigma_B = \sum_{c=1}^{2} (\mu_c - \mu)(\mu_c - \mu)^T$$

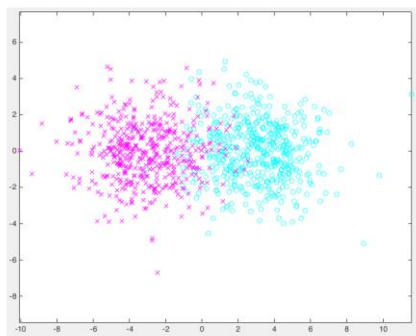
ب. تبدیل مناسبی را که ماتریس کواریانس عمومی Σ و ماتریس کواریانس بین کلاسی Σ_B را همزمان قطری می‌کند پیدا کنید. با اعمال این تبدیل روی داده‌ها ابعاد آنها کاهش می‌یابد (تصویر دوم).

ج. ماتریس‌های مذکور بعد از اعمال این تبدیل به چه شکلی در می‌آیند؟

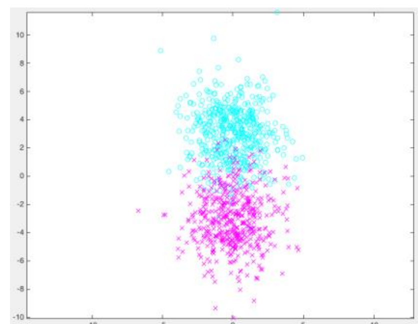
د. تصویر سوم نیز پس از اعمال تبدیل مشابه به دست آمده است. توضیح دهید که تفاوت موجود در تصویر دوم و سوم چگونه اتفاق افتاده است.



تصویر اول



تصویر دوم



تصویر سوم

۲- (اختیاری) آیا ماتریس $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ را می‌توان قطری کرد؟ آیا این ماتریس را می‌توان معرف پراکندگی داده‌های یک کلاس با توزیع نرمال دانست؟

۳- (اختیاری) آیا تبدیل سفیدکنندگی ارتونرمال است؟ چرا؟

۴- (اختیاری) یکی از کاربردهای قطری کردن ماتریس محاسبه ساده‌تر توانی از آن است. ماتریس A با مقادیر ویژه ۱و ۲ و بردارهای ویژه $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ و $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ را در نظر بگیرید.

الف. ماتریس A را بیابید.

ب. ماتریس A^3 را محاسبه کنید.

ج. ماتریس $B = \Phi \Lambda^3 \Phi^{-1}$ را که در آن Λ ماتریس مقادیر ویژه و Φ ماتریس بردارهای ویژه A هستند را بیابید و با ماتریس حاصل در قسمت ب مقایسه کنید.

د. ثابت کنید $A^n = \Phi\Lambda^n\Phi^{-1}$.

۵- (اختیاری) فرض کنید S یک ماتریس $n\times n$ بوجود آمده از دو بردار $\mathbf{m}_1$ و $\mathbf{m}_2$ بصورت $S=\mathbf{m}_1\mathbf{m}_1^T+\mathbf{m}_2\mathbf{m}_2^T$ می‌باشد. طول $\mathbf{m}_1$ و $\mathbf{m}_2$ بترتیب ۱ و ۲ و زاویه بین آنها 60° است. مقادیر ویژه S را بدست آورید.

۶. (سوال امتحان میان ترم بازشناسی آماری الگو سال ۱۳۹۳) اگر $\mathbf{x}$ یک بردار تصادفی در فضای n بعدی با متوسط صفر و ماتریس کواریانس Σ باشد، در حالت کلی می‌توان یک تبدیل متعامد یکه (ارتونرمال) با ماتریس Φ بدست آورد و متغیر n بعدی $\mathbf{z}$ را بصورت زیر تعریف کرد:

$$\mathbf{z} = \Phi^t\mathbf{x}$$

در رابطه فوق ماتریس تبدیل Φ متشکل از بردارهای ویژه ماتریس کواریانس Σ می‌باشد و داریم:

$$\Phi^t\Phi\Sigma = \Lambda$$

در رابطه اخیر، Λ یک ماتریس قطری متشکل از مقادیر ویژه Σ که به ترتیب نزولی مرتب شده‌اند می‌باشد.

الف- با توجه به اینکه برخی مقادیر ویژه بیش از حد کوچک بوده و کاهش ابعاد فضا دارای توجیه می‌باشد، ماتریس Φ' را بر حسب بردارهای ویژه Σ (یعنی $\boldsymbol{\varphi}_i$) طوری مشخص کنید که ابعاد فضای ویژگی را به n' ($n' < n$) کاهش دهد و متغیر n' بعدی حاصله ($\mathbf{z}'$) را در فضای جدید بر حسب $\mathbf{x}$ و Φ' بیان نمایید.

ب- نشان دهید امید ریاضی بردار (n' بعدی) $\mathbf{z}'$ برابر صفر است یعنی:

$$E\{\mathbf{z}'\} = \mathbf{0}$$

ج- نشان دهید واریانس مولفه i-ام در بردار $\mathbf{z}'$ ($i\leq n'$) برابر λ_i (عنصر i-ام قطر اصلی ماتریس Λ یا همان i-امین مقدار ویژه Σ) می‌باشد. یعنی نشان دهید:

$$E\{(z_i')^2\} = \lambda_i$$

(راهنمایی: $z_i' = \boldsymbol{\varphi}_i^t\mathbf{x}$)

د- با در اختیار داشتن $\mathbf{z}'$ و ماتریس‌های Φ و Φ' روابط مورد استفاده برای بازسازی بردار اولیه در فضای n بعدی (که اینبار آن را با $\mathbf{x}'$ نشان می‌دهیم) بنویسید.

(راهنمایی: با افزودن عناصر صفر به انتهای بردار $\mathbf{z}'$ ابعاد آن را از n' به n برگردانید و بردار حاصله را $\mathbf{z}_e'$ بنامید. سپس با انجام تبدیل وارون روی $\mathbf{z}_e'$ بردار اولیه $\mathbf{x}'$ را بازسازی نمایید.)

پاسخ:

$$\mathbf{z}_e' = \begin{bmatrix}\mathbf{z}'\\ \mathbf{0}\end{bmatrix}$$

$$\mathbf{x}' = \Phi\mathbf{z}_e'$$

ه- خطای مربع متوسط ایجاد شده در اثر کاهش ابعاد فضا و رسیدن به بردار (n' بعدی) $\mathbf{z}'$ و سپس بازگشت به فضای n بعدی اولیه و رسیدن به بردار بازسازی شده $\mathbf{x}'$ بصورت زیر تعریف می‌گردد:

$$e_{ms} = E\{\|\mathbf{x}-\mathbf{x}'\|^2\}$$

نشان دهید که خطای فوق برابر است با:

$$e_{ms} = \sum_{i=n'+1}^{n}\lambda_i$$

که در آن λ_i ها مقادیر ویژه ماتریس کواریانس Σ متناظر با ابعاد فضای حذف شده هستند.

پاسخ:

$$e_{ms} = E\{\|\mathbf{x}-\mathbf{x}'\|^2\} = E\{(\mathbf{x}-\mathbf{x}')^t(\mathbf{x}-\mathbf{x}')\} = E\{(\Phi\mathbf{z}-\Phi\mathbf{z}_e')^t(\Phi\mathbf{z}-\Phi\mathbf{z}_e')\}$$

$\Phi^t \mathrm{E}\{(\mathbf{z}-\mathbf{z}'_e)^t(\mathbf{z}-\mathbf{z}'_e)\}\Phi=\sum_{i=n'+1}^{n}\lambda_i$

و- با توجه به خطای بدست آمده در بند قبل یک معیار برای انتخاب n' (ابعاد فضای کاهش یافته) بیان کنید.

۷- (اختیاری) (سوال امتحانی میان ترم بازشناسی آماری الگو سال ۱۳۸۶)

در فضای n بعدی فرض کنید N نمونه در اختیار است: $X=\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$

می خواهیم مساله تحلیل مولفه‌های اصلی را در مورد این نمونه‌ها بررسی کنیم (شکل م-۶).

الف. بردار $\mathbf{m}$ را بر حسب نمونه‌های مجموعه X طوری بیابید که مجموع مربعات فاصله (اقلیدسی) آن از همه نمونه‌ها حداقل باشد.

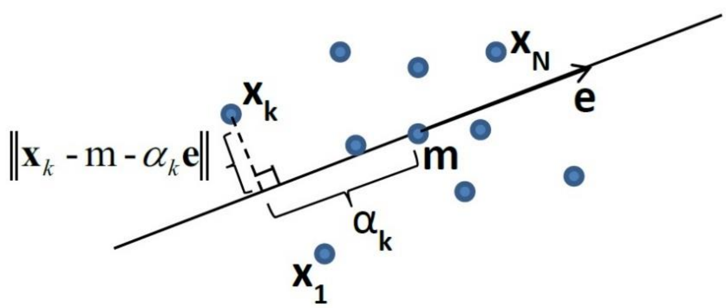
ب. حال فرض کنید که از نقطه $\mathbf{m}$ بدست آمده در فرض الف خطی در راستای بردار دلخواه $\mathbf{e}$ به طول واحد عبور کند. هر نقطه روی این خط در معادله زیر صدق می‌کند:

$\mathbf{x}=\mathbf{m}+\alpha\mathbf{e}$

که در آن ضریب α فاصله نقطه $\mathbf{x}$ از نقطه $\mathbf{m}$ را مشخص می‌کند. مجموعه ضرایب $\{\alpha_1, \alpha_2, \dots, \alpha_N\}$ را طوری بدست آورید که مجموع مربعات فاصله (اقلیدسی) بین نمونه‌های مجموعه X و بردارهای متناظر آنها روی خط (که با α_k مشخص می‌شوند) حداقل گردد.

ج. با استفاده از نتایج فرض‌های الف و ب نشان دهید راستای بهینه $\mathbf{e}$ که مجموع مربعات فواصل (اقلیدسی) بین نمونه‌ها و نقاط متناظر آنها روی خط گذرنده از نقطه $\mathbf{m}$ در راستای $\mathbf{e}$ را حداقل می‌سازد همان بردار ویژه متناظر با بزرگترین مقدار ویژه تخمین ماتریس کواریانس می‌باشد:

$\hat{S}=\frac{1}{N}\sum_{k=1}^{N}(\mathbf{x}_k-\mathbf{m})(\mathbf{x}_k-\mathbf{m})^{\mathrm{T}}$



شکل م-۶

۸- (اختیاری) روابط زیر برحسب $\mathbf{m}_1$، $\mathbf{m}_2$، Σ_1، Σ_2، $P(\omega_1)$ و $P(\omega_2)$ در بین دو توزیع برقرار است:

$P(\omega_1)(\Sigma_1+\mathbf{m}_1\mathbf{m}_1^{\mathrm{T}})+P(\omega_2)(\Sigma_2+\mathbf{m}_2\mathbf{m}_2^{\mathrm{T}})=I$

$P(\omega_1)\mathbf{m}_1+P(\omega_2)\mathbf{m}_2=0$

مطلوب است محاسبه کمیت‌های زیر بر حسب $\mathbf{m}_1$، $P(\omega_1)$ و $P(\omega_2)$:

(الف) $[P(\omega_1)\Sigma_1+P(\omega_2)\Sigma_2]^{-1}$

(ب) $(\mathbf{m}_2-\mathbf{m}_1)^{\mathrm{T}}[P(\omega_1)\Sigma_1+P(\omega_2)\Sigma_2]^{-1}(\mathbf{m}_2-\mathbf{m}_1)$

(ج) $|P(\omega_1)\Sigma_1+P(\omega_2)\Sigma_2|$

۹. (سوال امتحان میان ترم بازشناسی آماری الگو سال ۱۳۸۹)

اگر در فضای n بعدی، N نمونه آموزشی از M کلاس مختلف به طریقی در اختیار باشد که N_1 نمونه به کلاس ω_1، N_2 نمونه به کلاس ω_2، .. و N_M نمونه به کلاس ω_M تعلق داشته و داشته باشیم:

$$N = \sum_{j=1}^{M} N_j$$

مرکز ثقل عمومی مجموعه نقاط آموزشی فوق به صورت زیر تعریف می‌گردد:

$$\widehat{\boldsymbol{m}} = \frac{1}{N}\sum_{j=1}^{N} \boldsymbol{x}_j$$

و مرکز ثقل هر کلاس ω_i به صورت زیر تعریف می‌گردد:

$$\widehat{\boldsymbol{m}}_i = \frac{1}{N_i}\sum_{j=1}^{N_i} \boldsymbol{x}_{ij}$$

که در آن منظور از $\boldsymbol{x}_{ij}$ نمونه‌های متعلق به کلاس $\boldsymbol{i}$ام می‌باشد.

الف. $\widehat{\boldsymbol{m}}$را بر حسب$\widehat{\boldsymbol{m}}_i$محاسبه نمایید.

ب. اگر $\hat{\Sigma}_B$ماتریس کواریانس بین کلاس‌ها و $\hat{\Sigma}_W$ماتریس کواریانس درون کلاس‌ها و$\hat{\Sigma}$ماتریس کواریانس عمومی بوده و به صورت زیر تعریف گردند:

$$\hat{\Sigma}_B = \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{N_i}(\widehat{\boldsymbol{m}}_i - \widehat{\boldsymbol{m}})(\widehat{\boldsymbol{m}}_i - \widehat{\boldsymbol{m}})^T$$

$$\hat{\Sigma}_W = \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{N_i}(\boldsymbol{x}_{ij} - \widehat{\boldsymbol{m}}_i)(\boldsymbol{x}_{ij} - \widehat{\boldsymbol{m}}_i)^T$$

$$\hat{\Sigma} = \frac{1}{N}\sum_{i=1}^{M}\sum_{j=1}^{N_i}(\boldsymbol{x}_{ij} - \widehat{\boldsymbol{m}})(\boldsymbol{x}_{ij} - \widehat{\boldsymbol{m}})^T$$

نشان دهید :

$$\hat{\Sigma} = \hat{\Sigma}_B + \hat{\Sigma}_W$$

ج. چنانچه متغیر جدیدی را با استفاده از یک بردار n بعدی $\boldsymbol{a}$ به صورت زیر تعریف کنیم:

$$Z_i = \boldsymbol{a}^T(\boldsymbol{x}_i - \widehat{\boldsymbol{m}})$$

واریانس متغیر Z_i را محاسبه و آن را بر حسب $\hat{\Sigma}$ بیان نمایید.

د. می‌خواهیم بردار $\boldsymbol{a}$ را که کمیت $\frac{\boldsymbol{a}^T\hat{\Sigma}_B\boldsymbol{a}}{\boldsymbol{a}^T\hat{\Sigma}_W\boldsymbol{a}}$ را ماکزیمم نماید بیابیم. ماکزیمم کردن کمیت فوق به چه معنی است و به چه صورت در دسته بندی مفید می‌باشد؟

ه. نشان دهید که ماکزیمم کردن $\frac{\boldsymbol{a}^T\hat{\Sigma}_B\boldsymbol{a}}{\boldsymbol{a}^T\hat{\Sigma}_W\boldsymbol{a}}$ معادل با ماکزیمم کردن $\frac{\boldsymbol{a}^T\hat{\Sigma}_B\boldsymbol{a}}{\boldsymbol{a}^T\hat{\Sigma}\,\boldsymbol{a}}$ می‌باشد.

و. ماکزیمم کردن $\frac{\boldsymbol{a}^T\hat{\Sigma}_B\boldsymbol{a}}{\boldsymbol{a}^T\hat{\Sigma}\,\boldsymbol{a}}$ معادل است با ماکزیمم کردن $\boldsymbol{a}^T\hat{\Sigma}_B\,\boldsymbol{a}$با محدودیت $\boldsymbol{a}^T\hat{\Sigma}\,\boldsymbol{a} = 1$. با استفاده از روش لاگرانژ کسر فوق را با محدودیت بیان شده ماکزیمم نمایید و با استفاده از آن رابطه بردار $\boldsymbol{a}$را با ماتریس‌های$\hat{\Sigma}_B$و $\hat{\Sigma}$ به دست آورید. چه نتیجه‌ای از بحث فوق می‌توان گرفت؟

۱۰. (اختیاری) ماتریسهای کواریانس زیر از دو توزیع در اختیار هستند. تبدیلی که قطری‌سازی همزمان دو ماتریس را انجام دهد بدست آورید.

$$\Sigma_1 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}, \qquad \Sigma_2 = \begin{bmatrix} 1+\sqrt{3}/4 & 0.5 \\ 0.5 & 1-\sqrt{3}/4 \end{bmatrix}$$

۱۱- (اختیاری) نشان دهید بردارهای $\mathbf{S}^{-1}\mathbf{m}$ و $\Sigma^{-1}\mathbf{m}$ بردارهای یکسان با طول متفاوت هستند.

۱۲- (اختیاری) مقدار ویژه غیر صفر و بردار ویژه متناظر $\Sigma^{-1}\mathbf{m}\mathbf{m}^{\mathbf{T}}$ را برحسب Σ و $\mathbf{m}$ بدست آورید (راهنمایی: ماتریس $\Sigma^{-1}\mathbf{m}\mathbf{m}^{\mathbf{T}}$ از مرتبه واحد است).

۱۳- (اختیاری) سوال امتحانی میان ترم ۱۳۹۲

در زیر بحث تحلیل مولفه‌های اصلی (PCA) را با اندک تفاوت در روشی که در کلاس گفته شد مطرح می‌کنیم. اینبار بردار n بعدی $\boldsymbol{\varphi}$ بطول واحد را چنان جستجو می‌کنیم که کمیت زیر را حداکثر سازد:

$$J(\boldsymbol{\varphi}) = \boldsymbol{\varphi}^t \boldsymbol{\Sigma} \boldsymbol{\varphi}$$

که در آن $\boldsymbol{\Sigma}$ معرف ماتریس کواریانس است.

الف- به این منظور با استفاده از روش لاگرانژ تابع هزینه مساله فوق را تعریف کرده و با حداقل‌سازی آن نشان دهید که بردار $\boldsymbol{\varphi}$ مورد جستجو بردار ویژه ماتریس $\boldsymbol{\Sigma}$ متناظر با بزرگترین مقدار ویژه آن می‌باشد، یعنی داریم:

$$\boldsymbol{\Sigma}\boldsymbol{\varphi} = \lambda \boldsymbol{\varphi}$$

ب- روش بند الف یک راه حل تحلیلی را برای محاسبه بردار مطلوب $\boldsymbol{\varphi}$ معرفی می‌کند. در اینجا می‌خواهیم با استفاده از روش گرادیان نزولی، یک راه حل مرحله‌ای (iterative) برای محاسبه $\boldsymbol{\varphi}$ بدست آوریم. بطوریکه با شروع از یک حدس اولیه $\boldsymbol{\varphi}_0$ بتوان پس از طی چند مرحله، بردار مطلوب $\boldsymbol{\varphi}$ را محاسبه نمود. فرض کنید ماتریس $\boldsymbol{\Sigma}$ در اختیار باشد. با استفاده از تابع هزینه مناسب، این الگوریتم مرحله‌ای را که در آن $\boldsymbol{\varphi}_{k+1}$ فقط برحسب $\boldsymbol{\varphi}_k$، $\boldsymbol{\Sigma}$ و μ (نرخ آموزش) بیان شده باشد، بدست آورید.

راهنمایی: رابطه مربوط به بهینه‌سازی تابع هزینه بر اساس الگوریتم مرحله‌ای بصورت زیر است:

$$\boldsymbol{\varphi}_{k+1} = \boldsymbol{\varphi}_k - \mu \nabla \boldsymbol{J}(\boldsymbol{\varphi}_k)$$

ج- روش بند (ب) مستلزم در اختیار داشتن ماتریس $\boldsymbol{\Sigma}$ است و لذا گفته می‌شود که بصورت دسته‌ای (batch mode) عمل می‌کند. چنانچه $\boldsymbol{\Sigma}$ در اختیار نباشد، و فقط تعداد N نمونه از یک کلاس در اختیار باشد:

$$X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$$

می‌خواهیم یک راه حل برخط (online) برای تخمین مرحله‌ای بردار مورد جستجو بدست آوریم. به این منظور کافی است در الگوریتم بدست آمده در بند قبل، تخمین مرحله k-ام ماتریس کواریانس ($\widehat{\boldsymbol{\Sigma}}_k$) را جایگزین $\boldsymbol{\Sigma}$ نماییم. رابطه زیر این تخمین را بر حسب بردار جدید $\mathbf{x}_k$، تخمین مرحله k-ام میانگین ($\mathbf{m}_k$) و تخمین مرحله (k-1)-ام ماتریس کواریانس ($\widehat{\boldsymbol{\Sigma}}_{k-1}$) بدست می‌دهد:

$$\widehat{\boldsymbol{\Sigma}}_k = (1-(k-1)/k)(\mathbf{x}_k - \widehat{\mathbf{m}}_k)(\mathbf{x}_k - \widehat{\mathbf{m}}_k)^t + ((k-1)/k)\,\widehat{\boldsymbol{\Sigma}}_{k-1}$$

برنامه ساده‌ای (مثلا تحت متلب) بنویسید که تخمین بردار ویژه مطلوب $\boldsymbol{\varphi}$ را بصورت برخط و با استفاده از بردارهای ورودی $\mathbf{x}_k$ (k=1,2,…,N) محاسبه می‌کند. توجه شود که در هر مرحله از اجرای برنامه شما $\boldsymbol{\varphi}_{k+1}$ برحسب k، $\mathbf{x}_k$، $\boldsymbol{\varphi}_k$، $\widehat{\boldsymbol{\Sigma}}_k$ ، $\widehat{\mathbf{m}}_k$ و μ (نرخ آموزش) محاسبه می‌شود.

۱۴. (اختیاری) نشان دهید که در PCA، انتخاب K بردار ویژه متناظر با K مقدار ویژه بزرگتر ماتریس Σ، معادل انتخاب K بردار ویژه متناظر با K بردار ویژه کوچکتر ماتریس Σ^{-1} است.

۱۵. (اختیاری) نشان دهید که اگر ماتریس A و B همزمان قطری شوند آنگاه : $AB = BA$

(راهنمایی: دو ماتریس A و B همزمان قطری می‌شوند اگر ماتریس P وجود داشته باشد به شکلی که:

$D_A = P^{-1}AP$ و $D_B = P^{-1}BP$)

۱۶. (اختیاری) فرض کنید داده‌های نرمال X با میانگین μ و ماتریس کوواریانس Σ باشند. با مثالی به شکل پارامتری توضیح دهید که چرا اعمال ماتریس کواریانس داده‌ها روی آنها ($Y = \Sigma X$)، باعث چرخش توزیع به سمت بیشترین پراکندگی می‌شود؟ به توان رساندن این تبدیل(Σ^m) قبل از اعمال آن روی داده‌ها چه نتیجه‌ای دارد؟

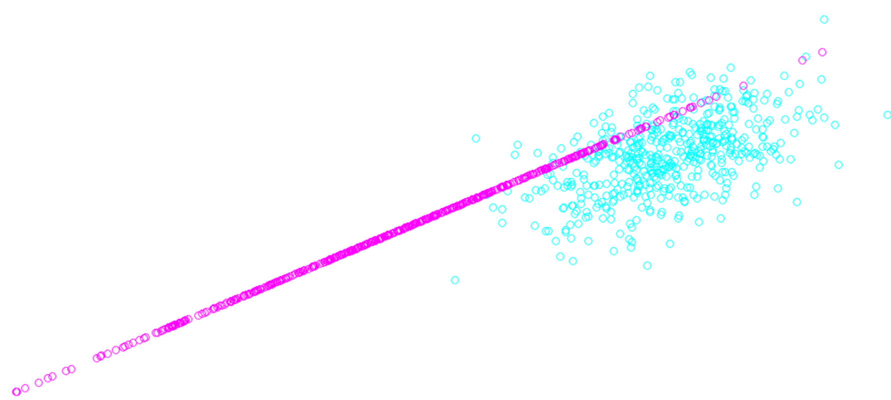
۱۷. دو تصویر زیر را در نظر بگیرید. تصویر اول داده‌های دو کلاس فیروزه‌ای و ارغوانی به ترتیب با ماتریس‌های کواریانس $\begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix}$ و $\begin{bmatrix} 4 & 6 \\ 6 & 9 \end{bmatrix}$ را نشان می‌دهد. یک تبدیل خطی روی داده‌های این دو کلاس اعمال شده که در نتیجه آن ماتریس کواریانس هر دو کلاس قطری شده‌اند (تصویر دوم).

الف. این تبدیل را بیابید.

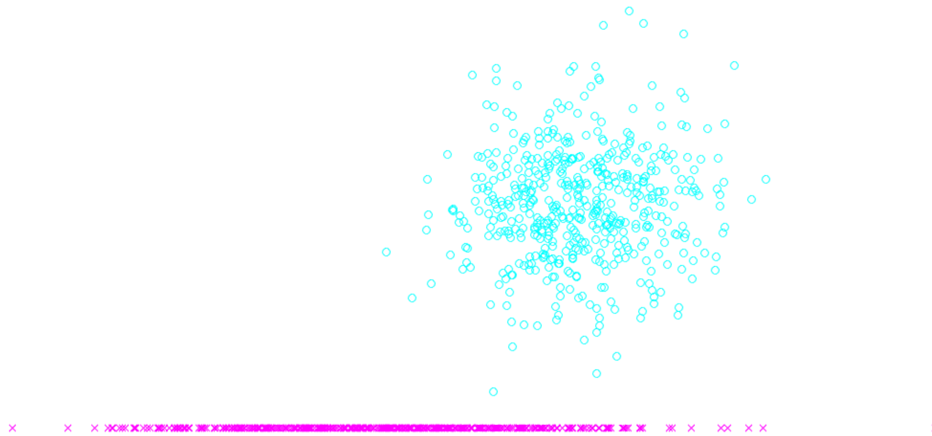
ب. ماتریس‌های کواریانس را بعد از اعمال این تبدیل محاسبه کنید.

ج. دسته‌بندی داده‌ها در کدام تصویر کارایی بیشتری دارد؟

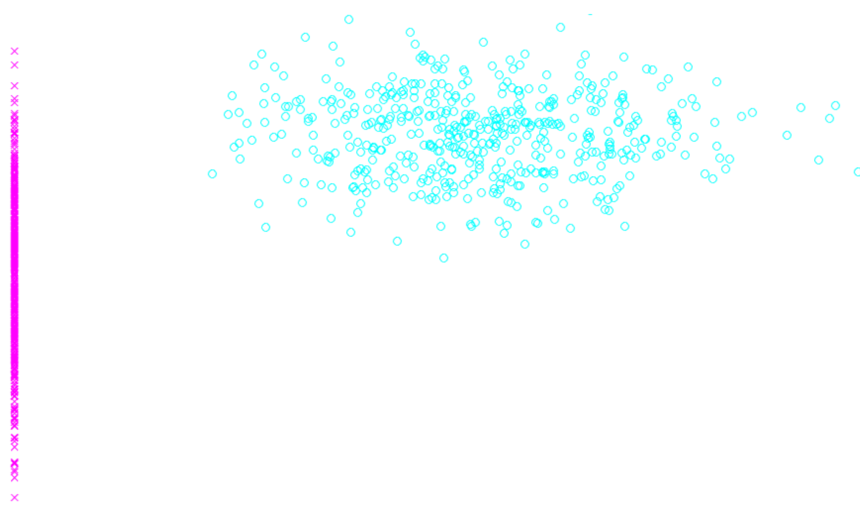
د. تصویر سوم نیز پس از اعمال تبدیل مشابه به دست آمده است. توضیح دهید که تفاوت موجود در تصویر دوم و سوم چگونه اتفاق افتاده است.



تصویر اول



تصویر دوم



تصویر سوم

۱۸- (میان ترم ۱۴۰۱) اگر یک ماتریس کوواریانس به صورت زیر داشته باشیم ، برای اینکه کمتر از ۳۰ درصد اطلاعات را از دست بدهیم ، حداقل باید چند بردار ویژه را در الگوریتم PCA داشته باشیم ؟

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 3 & 0 & 0 & 0 & 5 \end{bmatrix}$$

۱۹- میان ترم ۱۴۰۲

فرض کنید به کمک الگوریتم PCA متغیر تصادفی با ماتریس کواریانس $\begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 4 \end{bmatrix}$ را به فضای دو بعدی برده ایم.

الف)به طور متوسط چند درصد از اطلاعات پس از اعمال این تبدیل از بین میرود؟

ب) اگر مشاهده $x = \begin{bmatrix}1\\1\\1\end{bmatrix}$ از این متغیر تصادفی، مطابق تبدیل ذکر شده، به فضای دو بعدی برده شود و سپس به فضای ویژگی اولیه بازگردانده شود، خطای بازیابی(فاصله اقلیدسی بردار اولیه و بردار بازیابی شده) چقدر است؟

۲۰- میان ترم ۱۴۰۳

با فرض اینکه مقادیر ویژه ماتریس کواریانس داده‌ها غیر صفر و نابرابر باشند، کدامیک از عبارات زیر در مورد PCA درست است؟ <u>برای انتخاب خود دلیل ذکر کنید.</u>

الف- افزایش یک بعد با مقدار ثابت ۱ به انتهای همه بردارهای ویژگی، نتیجه PCA را تغییر نمی‌دهد (جز اینکه یک مقدار ویژه اضافی صفر به مقادیر ویژه ماتریس کواریانس اضافه می شود و اینکه بردارهای ویژه مفید یک بعد اضافی با مقدار صفر در انتهای خود خواهند داشت)

ب- اگر از PCA برای کاهش بعد از فضای d یعدی اولیه به فضای d′ بعدی و مجددا از PCA برای کاهش بعد از فضای d′ بعدی به فضای d′′ بعدی (d>d′>d′′) استفاده کنیم، نتیجه مشابه آن است که مستقیما از PCA برای کاهش بعد از فضای d بعدی اولیه به فضای d′′ بعدی استفاده کنیم.

ج- اگر همه بردارهای ویژگی، یک دوران ثابت دلخواه را قبل از اعمال PCA تجربه کنند، راستاهای متناظر با مولفه های اصلی تغییری نمی‌کنند.

د- اگر همه بردارهای ویژگی، یک دوران ثابت دلخواه را قبل از اعمال PCA تجربه کنند، بزرگترین مقدار ویژه ماتریس کواریانس نمونه ها ثابت می‌ماند.

موفق باشید