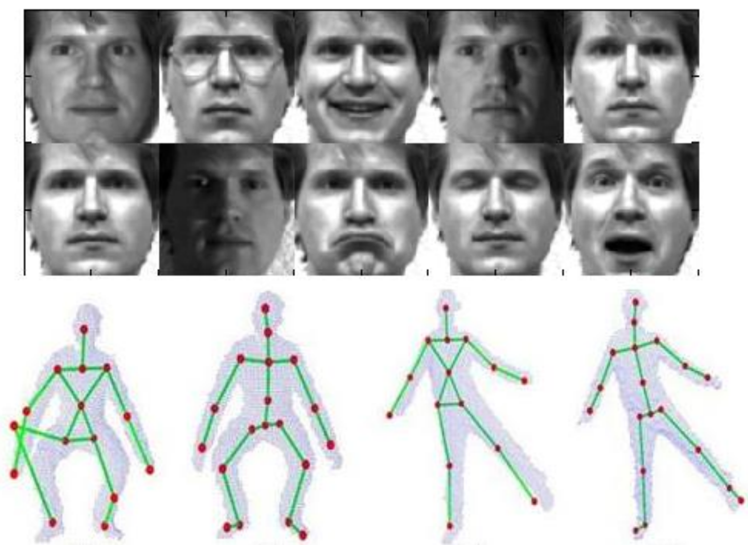
بسمه تعالی

دانشگاه صنعتی شریف

دانشکده مهندسی برق

گروه سیستم‌های دیجیتال



آزمایشگاه یادگیری و بینایی ماشین

دستور کار آزمایش اول: تحلیل مؤلفه های اساسی

زمان لازم برای انجام آزمایش: حداکثر دو جلسه

$$var = \sum_{i=1}^{3} [(a_i - m)^T u_1]^2 =$$

$$= u_1^T [(a_1 - m)(a_1 - m)^T + (a_2 - m)(a_2 - m)^T + (a_3 - m)(a_3 - m)^T] u_1 =$$

$$= u_1^T C u_1$$

در رابطه بالا، ماتریس C همان ماتریس کواریانس است. اگر مساله بالا را با شرط این که اندازه بردار یکه u1 برابر با یک است، به صورت مساله بهینه‌سازی حل کنیم، خواهیم داشت:

$$Cu = \lambda u$$

(امتیازی): رابطه بالا را اثبات نمایید. (راهنمایی: مساله بهینه‌سازی بالا یک مساله ضرب‌کننده لاگرانژ است؛ به این صورت که شما می‌خواهید واریانس را که در بالا بدست آمد، حداکثر کنید تا زیرفضا در جهت حداکثر پراکندگی قرار گیرد و در عین حال، اندازه $(u_1^T u_1 - 1)$ حداقل شود تا اندازه u1 که برابر با $u_1^T u_1$ است، برابر با یک شود.)

در رابطه بالا، λ ها همان مقادیر ویژه ماتریس کواریانس، و u هم بردارهای ویژه ماتریس کواریانس هستند. در مثال ما که دو بعد داریم، دو مقدار و بردار ویژه موجود است که مقدار ویژه‌ی متناظر با $u1$ از دیگری، به مراتب بزرگتر است؛ به طوریکه اگر از بین K (در این مثال، ۲) تا بعد، K' (در این مثال، ۱) تا مقدار ویژه را فقط داشته باشیم، داریم:

$$\frac{\sum_{j=1}^{K'} \lambda_j}{\sum_{j=1}^{K} \lambda_j} = \frac{\lambda_1}{\lambda_1 + \lambda_2} = 0.99$$

به مقادیر ویژه استخراج شده، اصطلاحا Principle Components گفته می‌شود. و به برداشتن مقادیر ویژه برتر و دور ریختن بقیه مقادیر ویژه Principle Component Analysis (PCA) گفته می‌شود.

فرض کنید دوباره بخواهیم داده‌های دوبعدی را بازسازی (reconstruct) کنیم. مثلا برای بازسازی a1 داریم:

$$\alpha_1 u_1 + m = \tilde{a}_1$$

دقت نمایید که اگر K' بعد را نگاه می‌داشتیم، بازسازی به صورت زیر می‌شد:

Bonus:

$$\|u\| = 1 = u^T u = 1$$

$$L(u,\lambda) = f(u) - \lambda \underline{(g(u) - c)}$$

$$\frac{\partial L(u,\lambda)}{\partial u} = 0 \quad , \quad \frac{\partial L(u,\lambda)}{\partial \lambda} = 0$$

$$L(u,\lambda) = u^T C u - \lambda(u^T u - 1)$$

$$\begin{cases} \dfrac{\partial L}{\partial u} = 2u^T C - 2\lambda u^T = 0 \Rightarrow u^T C = \lambda u^T \\[2ex] \dfrac{\partial L}{\partial \lambda} = u^T u - 1 = 0 \Rightarrow u^T u = 1 \end{cases}$$

$$\Rightarrow Cu = \lambda u$$

پیش گزارش

سه نقطهٔ (۱، ۱) و (۳، ۲) و(۲، ۳) را در یک فضای دوبعدی داریم.

- اگر بخواهیم این داده‌ها را به فضـای یک بعدی ببریم، مقدار ویژهٔ برتر و بردار ویژهٔ[1] نظیر آن را به دسـت آورید.
- داده‌ها را به فضای یک بعدی ببرید و خطای بازسازی را حساب کنید.

$$X_{raw} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 3 & 2 \end{bmatrix} , \quad \mu_1 = 2 \quad \mu_2 = 2$$

$$X_{centerized} = \begin{bmatrix} -1 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$C = X^T_{centerized} X_{centerized} = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$\Rightarrow$ Calculate the eigenvector & Value

[1] First Principle Component

$\Rightarrow C u_1 = \lambda_1 u_1 \Rightarrow (u_1, \lambda_1)? \Rightarrow$ Python!

$\underline{\lambda_1 = 3}$, $\underline{u_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}}$

$$X_{\text{centerized}} \underbrace{\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}}_{PC1 = u_1} = \frac{1}{\sqrt{2}} \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

reconstruction

$$XV = Z$$

$$\hat{X} = ZV^T \rightarrow X_{\text{recon}} = \hat{X} + \text{Mean}$$

$$\frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -2 & -2 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

$$\xrightarrow{+\text{Mean}} X_{\text{reconstructed}} = \begin{bmatrix} 1 & 1 \\ 2.5 & 2.5 \\ 2.5 & 2.5 \end{bmatrix}$$

$\mu_1 = 2$

$\mu_2 = 2$

Reconstruction Error $= \| X - X_{recon} \|_F$

$$= \left\| \begin{bmatrix} 0 & 0 \\ -0.5 & 0.5 \\ 0.5 & -0.5 \end{bmatrix} \right\|_F = 1$$