Principal Component Analysis

$\mathcal{D}, \{x^{(1)}, \ldots, x^{(n)}\}$

$x^{(i)} \in \mathbb{R}^d$

$X \in \mathbb{R}^d \quad X \sim P$

$X^{(i)} \sim P \quad$ i.i.d.

$E(X) = \mu$

$E(X-\mu)(X-\mu)^T = C$

$$a^* = \arg\min_{a} E(\|X - a\|^2) = \mu$$

$$\tilde{X} = Y u \qquad u \in \mathbb{R}^d, \quad \|u\| = 1$$

Let

$\mu = 0 \rightarrow E(X) = 0$

$$E(XX^T) = C$$

$$Y^* = \arg\min_{Y} \|X - Y u\|^2$$

$$= \arg\min_{Y} \|X\|^2 - 2 Y u^T X + Y^2$$

$$Y^* = u^T X$$

$$\hat{x} = (x^T u)\, u$$

$$E\left(\|x - \hat{x}\|^2\right)$$

$$= E\left(\|x\|^2\right) - 2E\left(x^T \hat{x}\right) + E\|\hat{x}\|^2$$

$$= E\left(\|x\|^2\right) - 2u^T C u + u^T C u$$

$$= \mathrm{Tr}(C) - u^T C u$$

$$\min_{u} E(\|X - \hat{X}\|^2)$$

$$= \min_{u} Tr(C) - u^T C u$$

$$\Rightarrow \max_{\substack{u \\ \|u\|=1}} u^T C u$$

Use Cauchy Schwartz

$$u^T v \leq \|u\| \, \|v\|$$

$$v = \lambda u$$

$$Cu = \lambda u$$

$$u^T v = \lambda \|u\|^2 = \lambda, \quad v = \lambda u$$

Largest value is λ_1.

$\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_d \geq 0$

$$C u_1 = \lambda_1 u_1.$$

$$u = u_1.$$

$$\hat{X} = (X^T u_1) u_1. \qquad Y_1 = X^T u_1$$
$$Y = X^T v$$

$$E \| X - Y_1 u_1 - Y v \|^2$$

$$= E \left(\| X - Y_1 u_1 \|^2 \right)$$
$$+ E \left(Y \| v \|^2 \right)$$
$$- 2 E \left(X - Y_1 u_1 \right)^T (Y v)$$

$$= E\left(\|X - Y_1 u_1\|^2\right)$$

$$- 2 E Y_1 (X^T v)$$

$$+ 2 E(Y_1 Y) u_1^T v$$

$$+ E(Y^2) \quad \left[E(Y Y_1) = 0 \right]$$

$$= E\left(\|X - Y_1 u_1\|^2\right)$$

$$- E(Y^2)$$

$$\min_{v} E\left(\|X - Y_1 u_1 - Y v\|^2\right)$$

$$= \min_{v} E\left(\|X - Y_1 u_1\|^2\right) - E\left(Y^2\right)$$

$$\arg\min_{v} E\left(\|X - Y_1 u_1 - Y v\|^2\right)$$

$$= \arg\min_{v} E\left(\|X - Y_1 u_1\|^2\right) - E\left(Y^2\right)$$

$$= \arg\max_{v} E\left(Y^2\right) = v^T C v$$

$$E(Y_1 Y) = 0 \Rightarrow u_1^T C v = 0$$

$$v = u_2 \Longrightarrow E(Y^2) = \lambda_2$$

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_n \end{bmatrix} = \begin{bmatrix} u_1^T X \\ u_2^T X \\ u_n^T X \end{bmatrix}$$

Canonical Component Analysis.

$$X \in \mathbb{R}^{d_1} \qquad Y \in \mathbb{R}^{d_2}$$

$$Z = \begin{bmatrix} X \\ Y \end{bmatrix} \qquad E(Z) = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}$$

$$E(X) = \mu_x \qquad E(Y) = \mu_Y$$

$$\operatorname{cov}(Z) = \begin{bmatrix} C_{xx} & C_{xy} \\ C_{yx} & C_{yy} \end{bmatrix}$$

$$C_{xy} = C_{yx}^T$$

$$\operatorname{cov}(Z) = E\left(Z - E(Z)\right)\left(Z - E(Z)\right)^T$$

$$R = U^T X \qquad S = V^T Y.$$

$$E(X) = 0, \quad E(Y) = 0$$

$$\rho_{RS} = \frac{E(RS)}{\sqrt{var(R)}\sqrt{var(S)}}$$

$$= \frac{U^T E(XY^T) V}{\sqrt{V^T E(YY^T) V}\sqrt{u^T E(XX^T) u}} = \rho_{uv}$$

$$E(YY^T) = C_{YY}$$

$$E(XX^T) = C_{XX}$$

$$E(XY^T) = C_{XY}$$

$$\rho(u, v) = \frac{u^T C_{XY} v}{\sqrt{v^T C_{YY} v}\sqrt{u^T C_{XX} u}}$$

$$\max_{u,v} \rho(u,v)$$

$\|u\| = 1$

$\|v\| = 1$

$$C_{xx}^{\frac{1}{2}} u = a$$

$$C_{yy}^{\frac{1}{2}} v = b$$

$$u = C_{xx}^{-1/2} a \quad , \quad v = C_{yy}^{-1/2} b$$

$$\rho(a,b) = \frac{a^T C_{xx}^{-1/2} C_{xy} C_{yy}^{-1/2} b}{\|a\| \quad \|b\|}$$

$$A = C_{xx}^{-1/2} C_{xy} C_{yy}^{-1/2}$$

$$\frac{a^T A b}{\|a\| \|b\|} \leq \frac{\|a\| \|Ab\|}{\|a\| \|b\|} = \frac{\|Ab\|}{\|b\|}.$$

$$\frac{b^T A^T A b}{\|b\|^2}$$ is maximized at

$$A^T A b = \lambda b$$

$$A^T A = C_{yy}^{-1/2} C_{yx} C_{xx}^{-1} C_{xy} C_{yy}^{-1/2}$$

$\lambda_1, \lambda_2 - - - , \lambda_{d_2}$

λ_1 is the largest eigenvalue of $A^T A$

g_1 be the eigenvector

$$\frac{a^T A b}{\|a\|\,\|b\|} \leq \frac{\|a\|\,\|Ab\|}{\|a\|\,\|b\|} = \frac{\|Ab\|}{\|b\|} = \sqrt{\lambda_1}$$

$A \in \mathbb{R}^{l \times m}$

$$A = \sum_{i=1}^{r} \sigma_i f_i g_i^T \qquad r = \min(l, m)$$

$f_i \in \mathbb{R}^l$, $g_i \in \mathbb{R}^m$

$$f_i^T f_j = 0 \quad i \neq j \qquad\Big|\qquad g_i^T g_j = 0 \quad i \neq j$$

$$ = 1 \quad i = j \qquad\Big|\qquad = 1$$

$j = 1, \ldots, r$

$$A g_i = \sigma_i f_i \qquad i = 1, \ldots, r$$

$$f_i^T A = \sigma_i g_i^T \qquad \Big|\qquad A^T f_i = \sigma_i g_i$$

$$\boxed{f_j^T A g_i = 0}$$

$$A^T A g_i = \sigma_i A^T f_i = \sigma_i^2 g_i$$

$$A A^T f_i = \sigma_i A g_i = \sigma_i^2 f_i$$

$b = g_1 \qquad v = C_{YY}^{-1/2} b = C_{YY}^{-1/2} g_1$

$Ab = \mu a$

(b, a) are singular vector pairs.

$A = C_{XX}^{-1/2} C_{XY} C_{YY}^{-1/2}$

$A = \sum_{i=1}^{r} \sigma_i f_i g_i^T$

$A g_i = \sigma_i f_i$

$a = f_1 \qquad u = C_{XX}^{-1/2} f_1$

Second pair is to choose

u, v such that

$u^T X$ and $u_1^T X$ are un-correlated

$v^T Y$ and $v_1^T Y$ are uncorrelated

$$\rho(u, v) = \frac{u^T C_{XY} v}{\sqrt{v^T C_{YY} v} \sqrt{u^T C_{XX} u}}$$

$$\max_{u, v} \rho(u, v)$$

$\|u\| = 1$, $\text{cov}(v^T Y, v_1^T Y) = 0$

$\|v\| = 1$, $\text{cov}(u^T X, u_1^T X) = 0$

$v^T C_{YY} v_1 = 0 \quad u^T C_{XX} u_1 = 0$

$$C_{xx}^{\frac{1}{2}} u = a \qquad\qquad a_1 = C_{xx}^{1/2} u_1$$

$$C_{yy}^{\frac{1}{2}} v = b \qquad\qquad b_1 = C_{yy}^{1/2} v_1$$

$$u^T C_{xx}^{1/2} C_{xx}^{1/2} u_1 = 0 \qquad\qquad v^T C_{yy}^{1/2} C_{yy}^{1/2} v_1 = 0$$

$$a^T a_1 = 0 \qquad\qquad b^T b_1 = 0$$

$$\rho(a,b) = \frac{a^T C_{xx}^{-1/2} C_{xy} C_{yy}^{-1/2} b}{\|a\| \; \|b\|}$$

$$a^T a_1 = 0 , \quad b^T b_1 = 0$$

$$\max_{a,b} \rho(a,b) \qquad\qquad a = f_2 , \; b = g_2$$

$$u = C_{xx}^{-1/2} f_2 , \quad v = C_{yy}^{-1/2} g_2$$

$$\tilde{X} = \begin{pmatrix} u_1^T X \\ \vdots \\ u_k^T X \end{pmatrix} \quad \tilde{Y} = \begin{pmatrix} v_1^T Y \\ \\ v_k^T Y \end{pmatrix}$$

$$\mathrm{cov}(u_l^T X, u_m^T X) \qquad \mathrm{cov}(v_l^T Y, v_m^T Y) = 0$$

$$= u_l^T C_{XX} u_m = 0 \qquad v_l^T C_{YY} v_m = 0$$

$$\mathrm{cov}(u_l^T X, v_m^T Y)$$

$$= u_l^T C_{XY} v_m$$

$$= f_l^T C_{XX}^{-1/2} C_{XY} C_{YY}^{-1/2} g_m \qquad l \neq m$$

$$= 0$$