Recap:

$$\bar{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} \sim f_{X_1 X_2 \ldots X_n}(x_1, \ldots x_n) \qquad \bar{\mu} \qquad K_{\bar{X}}$$

Assuming $\bar{X}$ is zero-mean:

$$K_{\bar{X}} = Q \Lambda Q^T$$

$\Lambda \in \mathbb{R}^{n \times n}$, diagonal.

$Q \in \mathbb{R}^{n \times n}$, orthonormal matrix, $Q^T Q = I$.

$$\Lambda = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

$\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n$.

$$Q = \begin{bmatrix} | & | & & | \\ \bar{q}_1 & \bar{q}_2 & \cdots & \bar{q}_n \\ | & | & & | \end{bmatrix}$$

$$\bar{y} = Q^T \bar{X}$$

$$Q\bar{y} = QQ^T \bar{X}$$

$$\bar{X} = Q\bar{y} = y_1 \bar{q}_1 + y_2 \bar{q}_2 + \ldots y_n \bar{q}_n.$$

$$\bar{X}_{app} = \sum_{i=1}^{m} y_i \bar{q}_i.$$

$$\mathbb{E}\left[\|\bar{e}\|^2\right] = \sum_{i=m+1}^{n} Var(y_i).$$

$$= \sum_{i=m+1}^{n} \lambda_i.$$

mean squared error.

If $\bar{X}$ is not zero mean, $\bar{X}$ has mean $\bar{\mu}$ and $K_{\bar{X}}$.

$\bar{X}_c = \bar{X} - \bar{\mu}$. has mean 0 and $K_{\bar{X}}$.

$$\bar{y} = Q\bar{X}_c$$

$$\bar{X}_c = Q\bar{y} = y_1 \bar{q}_1 + y_2 \bar{q}_2 + \ldots y_n \bar{q}_n$$

$$\bar{X}_c = \bar{\mu} + y_1 \bar{q}_1 + y_2 \bar{q}_2 + \ldots y_n \bar{q}_n$$

$$\bar{X}_{app} = \bar{\mu} + \sum_{i=1}^{m} y_i \bar{q}_i.$$

Principal Component Analysis:

Data: $\bar{X}^{(1)}, \bar{X}^{(2)}, \bar{X}^{(3)}, \ldots \bar{X}^{(5000)}$.

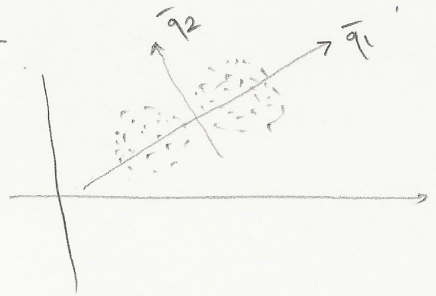
$\bar{X}^{(i)} \in \mathbb{R}^{1000}$.

e.g. $\bar{X}^{(i)}$: image of human faces

$\bar{X}^{(i)}$: transcription of 1000 genes in cell i.

Data visualution.

$n=2$:



Recipe:

1) Estimate the covariance matrix of the data and the its mean

$$\hat{\bar{\mu}} = \frac{1}{N}\sum_{i=1}^{N} \bar{X}_i \qquad \hat{K}_{\bar{X}}$$

2) Compute the eigendecomposition of $K_{\bar{X}}$

$$\hat{K}_X = Q \Lambda Q^T \qquad \bar{y} = Q(\bar{X} - \hat{\bar{\mu}}) \qquad y_i = \bar{q}_i^T(\bar{X} - \bar{\mu})$$

$$\bar{X}^{(i)} = y_1^{(i)} \bar{q}_1 + y_2^{(i)} \bar{q}_2 + \dots y_n^{(i)} \bar{q}_n + \hat{\bar{\mu}}$$

$\bar{q}_1$: first principal component.
(the eigenvector corresponding to the largest eigenvalue λ_1).
the direction of maximal variation

$\bar{q}_2$: second principal component.
the eigenvector corresponding to λ_2

Properties of $K_{\bar{X}}$.

1) $K_{\bar{X}}$ is real
2) $K_{\bar{X}}$ is symmetric

$$\left.\begin{matrix}\\ \\ \end{matrix}\right\} K_X = Q \Lambda Q^T \qquad \Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

3) $\lambda_1 \geq \lambda_2 \geq \dots \lambda_n \geq 0 \iff K_{\bar{X}}$ is positive semi-definit

$$\iff \bar{a}^T \underbrace{K_X}_{\in \mathbb{R}^{n \times n}} \bar{a} \geq 0 \quad \text{for every } \bar{a} \in \mathbb{R}^n.$$

$$\bar{a}^T \mathbb{E}\left[(\bar{X}-\bar{\mu})(\bar{X}-\bar{\mu})^T\right]\bar{a}$$

$$= \mathbb{E}\left[\underbrace{\bar{a}^T(\bar{X}-\bar{\mu})}_{z}\underbrace{(\bar{X}-\bar{\mu})^T\bar{a}}_{z^T = z}\right] = \mathbb{E}\left[z^2\right] \geq 0 .$$

Application: If $K_{\bar{X}}$ is positive semi-definite, then it has a "square-root", i.e $K_X = SS^T$

if $a \geq 0$, then $a = r^2$ for $a \in \mathbb{R}$.

$$\Lambda^{1/2} = \begin{bmatrix} \sqrt{\lambda_1} & & 0 \\ & \sqrt{\lambda_2} & \\ 0 & & \ddots \sqrt{\lambda_n} \end{bmatrix} \qquad K_{\bar{X}} = \underbrace{Q\Lambda^{1/2}}_{S}\underbrace{\Lambda^{1/2} Q^T}_{S^T}.$$

$$S^T = (Q\Lambda^{1/2})^T = (\Lambda^{1/2})^T Q^T$$

Coloring: $\bar{X}$ with $K_{\bar{X}} = I$, i.e. white.

We want to create $\bar{y}$ with covariance $K_{\bar{y}}$

$K_{\bar{y}} = SS^T$.

$\bar{y} = S\bar{X}$ $\quad K_{\bar{y}} = SS^T$

Whitening: $\bar{X}$ with covariance $K_{\bar{X}} = SS^T$.

$\bar{y} = S^{-1/2}\bar{X}$. assuming S is invertible

$$K_{\bar{y}} = S^{-1/2}\, \underbrace{SS^T (S^{-1/2})^T}_{(S^{-1/2}S)^T} = I .$$