Properties of K_x

1) real entries
2) symmetric

} by definition

eigenvalues, eigenvectors of $K \in \mathbb{R}^{n \times n}$

$\bar{v} \neq 0$ $\quad K\bar{v} = \lambda \bar{v}$, the $(\lambda, \bar{v})$ are called an (eigen value, eigenvector) pair for K.

$$(K - \lambda I)\,\bar{v} = 0$$

$$\Rightarrow \det(K - \lambda I) = 0$$

polynomial of degree n in λ called characteristic polynomial.

$\Rightarrow$ the n roots of this polynomial will be the eigenvalues of K.

$$a\lambda^2 + b\lambda + c = 0 \qquad \lambda = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

In general, the roots of the characteristic polynomial, i.e the eigenvalues, can be real or complex. However, if K is symmetric (and real) then the n eigenvalues are real.

Facts from linear algebra:

If K is real & symmetric, then

1). it has n eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_n$
and n eigenvectors $\bar{q}_1, \bar{q}_2, \ldots \bar{q}_n$

2) $\lambda_1, \lambda_2, \ldots \lambda_n$ are real

3) $\bar{q}_1, \bar{q}_2, \ldots \bar{q}_n$ can be chosen orthonormal, i.e. $\bar{q}_i^T \bar{q}_j = \begin{cases} 1 & i=j \\ 0 & o/w \end{cases}$

Let $Q = \begin{bmatrix} | & | & & | \\ \bar{q}_1 & \bar{q}_2 & \cdots & \bar{q}_n \\ | & | & & | \end{bmatrix}$ $\quad \Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$

orthonormal matrix $QQ^T = Q^TQ = I$

$$KQ = \begin{bmatrix} | & | & & | \\ \lambda_1 \bar{q}_1 & \lambda_2 \bar{q}_2 & \cdots & \lambda_n \bar{q}_n \\ | & | & & | \end{bmatrix}$$

$$= \begin{bmatrix} | & | & & | \\ \bar{q}_1 & \bar{q}_2 & \cdots & \bar{q}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \ \lambda_n \end{bmatrix} = Q\Lambda$$

$$KQ = Q\Lambda \qquad K\underbrace{QQ^T}_{I} = Q\Lambda Q^T \Rightarrow K = Q\Lambda Q^T$$

$\Rightarrow$ Fact: A real symmetric matrix can be factored into
$K = Q\Lambda Q^T$
~~w~~ where Λ is a diagonal
Q an orthonormal matrix, i.e. $QQ^T = Q^TQ = I$

(Spectral Decomposition)
Eigen Decomposition.

Recall: $\bar{y} = H\bar{X}$ $\quad K_{\bar{y}} = H K_{\bar{x}} H^T$

Application: $K_{\bar{X}} = Q \Lambda Q^T$ for some Q orthonormal, Λ diagonal.

$$H = Q^T \Rightarrow K_{\bar{y}} = Q^T K_{\bar{x}} Q = Q^T Q \Lambda Q^T Q = \Lambda$$

Λ diagonal.

$\bar{Y} = Q^T \bar{X}$ then components of $\bar{Y}$ are uncorrelated.

$$\mathrm{Cov}(y_j, y_i) = 0 \quad \text{if } i \neq j.$$

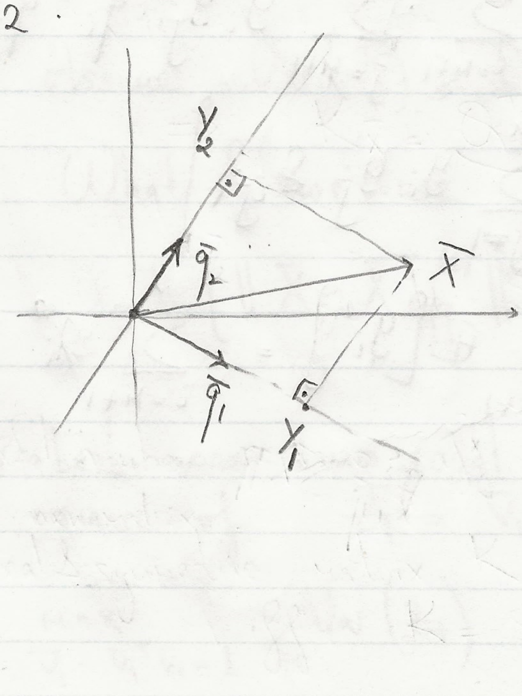
$$\mathrm{Var}(y_i) = \lambda_i \quad i = 1, \dots n.$$

$\bar{X} = Q\bar{Y}$ where $\bar{Y}$ is uncorrelated.

$$= \begin{bmatrix} | & | & & | \\ q_1 & q_2 & \cdots & q_2 \\ | & | & & | \end{bmatrix} \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} = Y_1 \bar{q}_1 + Y_2 \bar{q}_2 + \dots Y_n \bar{q}_n$$

uncorrelated with each other.

$n = 2$.



$Y_i = \bar{q}_i^T \bar{X}$

Karhunen-Loeve expansion of $\bar{X}$.

$$\mathrm{Var}(y_i) = \lambda_i$$

$$\underbrace{\sum_{i=1}^{n} \mathrm{Var}(X_i)}_{\mathrm{trace}(K_X)} = \underbrace{\sum_{i=1}^{n} \mathrm{Var}(y_i)}_{\lambda_1 + \dots + \lambda_n}$$

Assume $\bar{X}$ is zero-mean. $\mathbb{E}[\bar{X}] = 0$

and $\lambda_1 \geq \lambda_2 \geq \dots \lambda_n \geq 0$ are ordered.

$\Rightarrow$ $Var(y_1) \geq Var(y_2) \geq \dots Var(y_n) \geq 0$ are ordered.

If we want to have a simple (low-dimensional) approximation for $\bar{X}$, we can truncate the components with smaller variances

$$\bar{X}_{app} = \sum_{i=1}^{M} y_i \bar{q_i}$$

What happens when we use another orthonormal basis, i.e. orthonormal matrix V s.t $VV^T = V^TV = I$, to represent the vector X?

$$\bar{Z} = V^T \bar{X}$$

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

$$K_{\bar{Z}} = V^T K_X V$$

$$= \underbrace{V^T Q}_{\neq I} \Lambda \underbrace{Q^T V}_{\neq I}$$

$\bar{Z}$ does not have uncorrelated components because $K_{\bar{Z}}$ is not diagonal in general.

$$\bar{X} = V^T \bar{Z}$$

$$= \begin{bmatrix} | & | & & | \\ \bar{v}_1 & \bar{v}_2 & \cdots & \bar{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = z_1 \bar{v}_1 + z_2 \bar{v}_2 + \dots + z_n \bar{v}_n$$

not uncorrelated in general.

$$Var(z_i) = (K_Z)_{ii}$$

$$\sum_{i=1}^{n} Var(z_i) = Trace(K_{\bar{Z}}) = Trace(V^T K_X V) = Trace(K_X V V^T)$$

$$= Trace(K_X)$$

$$\Rightarrow \sum_{i=1}^{n} Var(z_i) = \sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} Var(X_i)$$

Assume $Var(z_1) \geq Var(z_2) \geq \dots Var(z_n) \geq 0$.

What if we use $\bar{X}_{app} = \sum_{i=1}^{m} z_i \bar{v}_i$

$$\bar{e} = \bar{X} - \bar{X}_{app}$$

$$\mathbb{E}\left[\|\bar{e}\|^2\right] = \mathbb{E}\left[\sum_{i=1}^{n} e_i^2\right].$$

appoximation error.

<u>Claim:</u> $\mathbb{E}\left[\|\bar{e}\|^2\right] = \sum_{i=m+1}^{n} Var(z_i)$.

<u>Goal:</u> In order toget the best possible m-dimensional representation for $\bar{X}$, choose V such that

$$\sum_{i=m+1}^{n} Var(z_i) \text{ is minimized.}$$

or equivalently $\sum_{i=1}^{m} Var(z_i)$ is maximized.

<u>Claim:</u> $\max_{V} \sum_{i=1}^{m} Var(z_i) = \sum_{i=1}^{m} \lambda_i$. where $\lambda_1 \geq \lambda_2 \dots \geq \lambda_n$

<u>Conclusion:</u> The KL expansion gives the best m-dimensional approximation for $\bar{X}$ in terms of squared l_2 error.