- 问题：选取变换矩阵 $\boldsymbol{\Phi}$，使得降维后的新向量在最小均方差条件下接近原来的向量 $\boldsymbol{x}$

对于 $\boldsymbol{x}=\sum_{j=1}^{n} a_j \boldsymbol{\varphi}_j$ ，现仅取 m 项，对略去的系数项用预先选定的常数 b 代替，此时对 $\boldsymbol{x}$ 的估计值为：

$$\hat{\boldsymbol{x}}=\sum_{j=1}^{m} a_j \boldsymbol{\varphi}_j+\sum_{j=m+1}^{n} b \boldsymbol{\varphi}_j$$

则产生的误差为：

$$\Delta \boldsymbol{x}=\boldsymbol{x}-\hat{\boldsymbol{x}}=\sum_{j=m+1}^{n}(a_j-b) \boldsymbol{\varphi}_j$$

则 $\Delta \boldsymbol{x}$ 的均方误差为：

$$\overline{\varepsilon^2}=E\{\|\Delta \boldsymbol{x}\|^2\}=\sum_{j=m+1}^{n}\{E(a_j-b)^2\}$$

要使 $\overline{\varepsilon^2}$ 最小，对 b 的选择应满足：

$$\frac{\partial}{\partial b}[E(a_j-b)^2]=\frac{\partial}{\partial b}[E(a_j^2-2a_j b+b^2)]=-2[E(a_j)-b]=0$$

因此，$b=\mathrm{E}[a_j]$，即对省略掉的 a 中的分量，应使用它们的数学期望来代替，此时的误差为：

$$\begin{aligned}\overline{\varepsilon^2}&=\sum_{j=m+1}^{n} E[(a_j-E\{a_j\})^2]=\sum_{j=m+1}^{n} \boldsymbol{\varphi}_j^T E[(\boldsymbol{x}-E\{\boldsymbol{x}\})(\boldsymbol{x}-E\{\boldsymbol{x}\})^T] \boldsymbol{\varphi}_j \\ &=\sum_{j=m+1}^{n} \boldsymbol{\varphi}_j^T \boldsymbol{C}_x \boldsymbol{\varphi}_j\end{aligned}$$

其中，$\boldsymbol{C}_{\mathrm{x}}$ 为 $\boldsymbol{x}$ 的协方差矩阵, $\{\boldsymbol{\varphi}_j\}$是正交向量，由拉格朗日法可导出 $\boldsymbol{\varphi}_j$ 为 $\boldsymbol{C}_{\mathrm{x}}$ 的特征值。

设 λ_j 为 $\boldsymbol{C}_{\mathrm{x}}$ 的第 j 个特征值，$\boldsymbol{\varphi}_j$ 是与 λ_j 对应的特征向量，则

$$\boldsymbol{C}_x \boldsymbol{\varphi}_j=\lambda_j \boldsymbol{\varphi}_j$$

由于

$$\boldsymbol{\varphi}_j^T \boldsymbol{\varphi}_j = 1$$

从而

$$\boldsymbol{\varphi}_j^T \boldsymbol{C}_x \boldsymbol{\varphi}_j = \lambda_j$$

因此

$$\overline{\varepsilon^2} = \sum_{j=m+1}^{n} \boldsymbol{\varphi}_j^T \boldsymbol{C}_x \boldsymbol{\varphi}_j = \sum_{j=m+1}^{n} \lambda_j$$

由此可以看出，λ_j 值越小，误差也越小。