

Input vector: $\mathbf{x} \in \mathbb{R}^D$

Transformation matrix: $\mathbf{A} = \begin{bmatrix} | & & | \\ \alpha_1 & \dots & \alpha_M \\ | & & | \end{bmatrix} \in \mathbb{R}^{D \times M}$

Finding 1st PCA dimension $\alpha_1 \in \mathbb{R}^D$

We would like to find $\alpha_1 \in \mathbb{R}^D$ that maximizes the variance:

$$\text{Var}\{\alpha_1^T \mathbf{x}\} = \alpha_1^T \Sigma \alpha_1$$

where Σ is the covariance of the data $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$.

Constrained optimization problem:

$$\max \alpha_1^T \Sigma \alpha_1, \quad \text{s.t.} \quad \alpha_1^T \alpha_1 = 1$$

Lagrange optimization:

$$\begin{aligned} L &= \alpha_1^T \Sigma \alpha_1 - \lambda(\alpha_1^T \alpha_1 - 1) \\ \Rightarrow \frac{\partial L}{\partial \alpha_1} &= 0 \Rightarrow \Sigma \alpha_1 - \lambda \alpha_1 = 0 \Rightarrow \lambda \alpha_1 = \Sigma \alpha_1 \end{aligned}$$

This is the eigenvector equation! We choose the eigenvector with the largest eigenvalue: $\Sigma \alpha_1$ is the projection of Σ onto the eigenvector α_1 . The largest this projection, the more variance vector α_1 will represent, therefore we have to choose the eigenvector with the largest eigenvalue.

Finding 2nd PCA dimension $\alpha_2 \in \mathbb{R}^D$

We would like to find $\alpha_2 \in \mathbb{R}^D$ that maximizes the variance $\text{Var}\{\alpha_2^T \mathbf{x}\}$, so that α_2 is *orthogonal* to α_1 .

Constrained optimization problem:

$$\max \alpha_2^T \Sigma \alpha_2, \quad \text{s.t.} \quad \alpha_2^T \alpha_2 = 1 \quad \text{and} \quad \alpha_2^T \alpha_1 = 0$$

Lagrange optimization:

$$\begin{aligned} L &= \alpha_2^T \Sigma \alpha_2 - \lambda(\alpha_2^T \alpha_2 - 1) - \phi \alpha_2^T \alpha_1 \\ \Rightarrow \frac{\partial L}{\partial \alpha_2} &= 0 \Rightarrow \Sigma \alpha_2 - \lambda \alpha_2 - \phi \alpha_1 = 0 \end{aligned}$$

If we right-multiply α_1 in the above expression:

$$\Sigma \alpha_2 \alpha_1^T - \lambda \alpha_2 \alpha_1^T - \phi \alpha_1 \alpha_1^T = 0 \Rightarrow \phi = 0$$

When $\phi = 0$, we get $\lambda \alpha_2 = \Sigma \alpha_2$. This corresponds to another eigenvalue equation and we choose the eigenvector with the second largest eigenvalue.