

Principal Components Analysis Notes

Alexander Yuan

September 13 2025

1 Background

1.1 History

Principal Components Analysis (PCA) goes all the way back to Harold Hotelling in the 1933 when the cutting edge topics in statistics at the time were the multivariate methods we know today. It was first introduced as a technique for obtaining a smaller set of orthogonal linear projections of a single collection of correlated variables where the projections are ordered by decreasing variance. Note that it is also called the *Karhunen-Loève transform* in communications theory and *empirical orthogonal functions* in atmospheric science.

1.2 Eigendecomposition

Let \mathbf{A} be a square n by n matrix with n linearly independent eigenvectors e_i for $i = 1, \dots, n$. Then \mathbf{A} can be factored as

$$\mathbf{A} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^{-1}$$

where \mathbf{E} is the square n by n matrix whose i 'th column is the eigenvector e_i of \mathbf{A} and $\mathbf{\Lambda}$ is the diagonal matrix whose diagonal elements are the corresponding eigenvalues i.e $\mathbf{\Lambda}_{i,i} = \lambda_i$. Note that this is a special case of the *spectral theorem*.

1.3 Singular Value Decomposition (SVD)

Let $T \in \mathcal{L}_{V \rightarrow W}$ where V and W are two finite dimensional inner product spaces both over \mathbb{C} . Let $T^* \in \mathcal{L}_{W \rightarrow V}$ be the *adjoint* of T and recall that a map is called *self-adjoint* if it equals its adjoint. From this, one can deduce that the eigenvalues of T^*T are non-negative.

Define the *singular values* of T to be the non-negative square roots of the eigenvalues of T^*T in decreasing order, each included as many times as the dimension of the corresponding eigenspace of T^*T . Some important properties of T^*T to be aware of is:

- i) T^*T is self-adjoint and $\langle T^*Tv, v \rangle \geq 0$ for all $v \in V$

- ii) $\text{null } T^*T = \text{null } T$
- iii) $\text{range } T^*T = \text{range } T^*$
- iv) $\dim \text{range } T = \dim \text{range } T^* = \dim \text{range } T^*T$

We can also characterize positive singular values:

- i) T is injective $\iff 0$ is not a singular value of T
- ii) the number of positive singular values of $T = \dim \text{range } T$
- iii) T is surjective \iff number of positive singular values of $T = \dim W$

Before we state the result, note that SVD is a fundamental result as it shows every linear map from V to W has a clean description in terms of its singular values and orthonormal lists in V and W . Techniques exist for approximating eigenvalues and eigenvectors of positive operators which makes SVD very useful.

Suppose $T \in \mathcal{L}_{V \rightarrow W}$ and let the positive singular values of T be s_1, s_2, \dots, s_m . Then there exists orthonormal lists e_1, \dots, e_m in V and f_1, \dots, f_m in W such that

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \dots + s_m \langle v, e_m \rangle f_m$$

for every $v \in V$

2 References

Axler, S. (2024). Linear Algebra Done Right (4th ed.). Springer.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd ed.). Springer

Ma, Y., & Fu, Y. (Eds.). (2011). Manifold Learning Theory and Applications. CRC Press.