Quick Recap Of Last Class

In the last class, our discussion was as follows:

Supposing we have N realizations of random variable vectors $x \in \mathbb{R}^n$ (random variable is a collection of all possible images of different kinds of dogs in Michigan state), the mean or expected value based on the N realizations, or assuming only N realizations exist

$$\mu = \sum_{i=1}^{N} P(X = \vec{x_i})\, x_i = \frac{1}{N} \sum_{i=1}^{N} x_i$$

The covariance matrix expresses the expectation of how any two components (or pixels) are related, i.e

$$\Sigma = \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)(x_i - \mu)^T$$

We showed that the eigen vectors of Σ are orthonormal to one another because it is PSD (and Hermitian); & the eigenvalues are real & non-negative.

We noted that this eigenvectors can serve a new coordinate system or basis for our random variable space; and more importantly the e.vectors are ordered in terms of relative importance of how much they express distinction or variance in the data.

We order the eigenvectors according to corresponding e.values

Note unbiased Covariance uses $\frac{1}{n-1}$

$$T = [v_1, v_2, \ldots, v_{n-1}]$$

Principal components

$$y = T^H x \leftarrow \text{PCA transformation}$$

which is the projection of the data onto the eigen vector space

Further $\Sigma = T \Lambda T^H$

Λ Diagonal Matrix of eigenvalues

Variance of Transformed Data

Notice that

$$\Sigma = \frac{1}{N}\sum_{i=1}^{N}(\vec{x}_i - \mu)(\vec{x}_i - \mu)^T$$

is equivalent to

$$\Sigma = \frac{1}{n}\left(X - \vec{\mu}\,\vec{1}^T\right)\left(X - \vec{\mu}\,\vec{1}^T\right)$$

where $X = \begin{bmatrix} | & | & & | \\ \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_N \\ | & | & & | \end{bmatrix}$ $\qquad \vec{1} = [\underbrace{1, 1, \ldots 1}_{n}]^T$

$$\Sigma = \frac{1}{n} X_m X_m^T$$

where X_m is the mean-subtracted data.

In the PCA-transformed space, the data has zero covariance

$$Y_m = T_n^H X_m$$

$$\Sigma_y = \frac{1}{n} Y_m Y_m^T = \frac{1}{n} T_n^H X_m X_m^H T_n$$

$$= T_n^H \Sigma_x T_n$$

But recall that $\Sigma_x = T_n \Lambda T_n^H$

$$\therefore \Sigma_y = T_n^H T_n \Lambda T_n^H T_n = \Lambda$$

Λ: Covariance is diagonal, so just Variance

Singular Value Decomposition

Every complex-valued matrix $A \in \mathbb{C}^{n \times m}$ can be decomposed as

$$A = USV^H$$

where $U \in \mathbb{C}^{n \times n}$ and $V \in \mathbb{C}^{m \times m}$ are unitary matrices with orthonormal columns, and $S \in \mathbb{R}^{n \times m}$ is a matrix with real non-negative entries on the diagonals and zeros off the diagonal

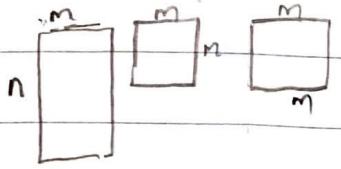
When $n \geq m$, the matrix S has at most m non-zero elements on the diagonal and may be written as

$$S = \begin{bmatrix} \hat{S} \\ 0 \end{bmatrix}$$

Therefore, we can exactly represent the matrix A using 'economy' SVD:

$$A = USV^H = [\hat{U} \; \hat{U}^{\perp}] \begin{bmatrix} \hat{S} \\ 0 \end{bmatrix} V^H = \hat{U}\hat{S}V^H$$

where the columns of $\hat{U}^{\perp}$ span a vector space that is complementary and orthogonal to that spanned by $\hat{U}$.



The columns of U are called left singular vectors; those of V are right singular vectors. The diagonal elements of $\hat{S}$ are called singular values and are ordered from largest to smallest The rank of matrix A is equal to the number of non-zero singular values

We can wrote the SVD for the mean subtracted data as

$$X_m = U S V^H$$

Recall that the covariance matrix is

$$\Sigma = \frac{1}{n} X_m X_m^H = \frac{1}{n} U S^2 U^H = T \Lambda T^H$$

Hence, the principal components of T are also the left singular vectors of X_m. The variance of the data in the new coordinates

$$\lambda_k = \frac{\sigma^2}{n}$$