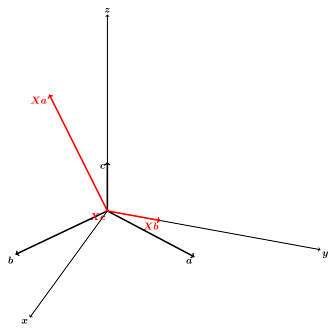
Internet Analytics (COM-308)

Homework Set 4

Exercise 1

In this exercise we study the geometric interpretation of singular value decomposition (SVD) seen in class.Assume the images of three vectors $a = [1, 2, 0]^\intercal, b = [2, -1, 0]^\intercal$ *and* $c = [0, 0, 1]^\intercal$ *under the linear transformation* $X = \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} \\ x_{2,1} & x_{2,2} & x_{2,3} \\ x_{3,1} & x_{3,2} & x_{3,3} \end{bmatrix}$ *is* $Xa = [3, 0, 4]^\intercal, Xb = [0, 1, 0]^\intercal$ *and* $Xc = [0, 0, 0]^\intercal$*, respectively.*



(a) Find the linear transformation X (matrix X) directly by solving the system of linear equations

$$X[a\ b\ c] = [Xa\ Xb\ Xc].$$

$$X = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 0.6 & 1.2 & 0 \\ 0.4 & -0.2 & 0 \\ 0.8 & 1.6 & 0 \end{bmatrix}$$

(b) The SVD decomposition of matrix X is

$$X = U\Sigma V^\intercal,$$

where we can interpret the three matrices U, Σ *and* $V^\intercal$ *as follows:(i) the two matrices* U *and* $V^\intercal$ *are changes of basis; and (ii) matrix* Σ *is scaling each axis (dimension) independently. Determine the transformation X from this interpretation of SVD (without any calculations).*

Note that the three vectors a, b and c are orthogonal, and can build new basis.

- Matrix $V^\intercal$ changes the xyz basis to the new abc basis:

$$v^\intercal = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- As $\frac{|Xa|}{a} = \sqrt{5}, \frac{|Xb|}{b} = \frac{1}{\sqrt{5}}$ and $\frac{|Xc|}{c} = 0$ we have

$$\Sigma = \begin{bmatrix} \sqrt{5} & 0 & 0 \\ 0 & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- Based on the computed $V^\intercal$ and Σ matrices we have

$$\Sigma v^\intercal [a\ b\ c] = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Therefore, with the following change of basis $U = \begin{bmatrix} 0.6 & 0 & -0.8 \\ 0 & 1 & 0 \\ 0.8 & 0 & 0.6 \end{bmatrix}$ we find the X transformation.

Note that since we have two non–zero singular values, for the third column of U we can use any other possible vector.

Exercise 2

Let X be an $n \times m$ matrix, where each of the n rows is a datum represented by an m-dimensional vector. Furthermore, assume that X is zero-mean, i.e. each column of X sums up to zero. Consider the singular value decomposition (SVD) of X:

$$X = U\Sigma V^\intercal.$$

(a) Express the SVD of the matrix $X^\intercal$ in terms of the SVD of X.

We have

$$X^\intercal = (U\Sigma V^\intercal)^\intercal = V\Sigma^\intercal U^\intercal.$$

(b) We define the $m \times m$ covariance matrix $\mathrm{Cov}_{m\times m}[X]$ as

$$\mathrm{Cov}_{m\times m}[X] = \begin{pmatrix} \mathrm{Cov}(X_1, X_1) & \mathrm{Cov}(X_1, X_2) & \dots & \\ \mathrm{Cov}(X_2, X_1) & \ddots & & \\ \vdots & & \mathrm{Cov}(X_m, X_m) \end{pmatrix}$$

where $\mathrm{Cov}(X_i, X_j) = \frac{1}{n}\sum_{k=1}^{n} X_{ki}X_{kj}$. Express $\mathrm{Cov}_{m\times m}[X]$ in terms of multiplication of two matrices.

We have

$$\mathrm{Cov}_{m\times m}[X] = \tfrac{1}{n}[X^\intercal X].$$

(c) Principal Component analysis (PCA) can be seen as the eigendecomposition of the covariance matrix:

$$\mathrm{Cov}_{m\times m}[X] = Q\Lambda Q^\intercal.$$

How does this interpretation of PCA relate to the SVD of X? Express Q and Λ in terms of U, Σ and V.

$$\begin{aligned} \mathrm{Cov}_{m\times m}[X] = \tfrac{1}{n}[X^\intercal X] &= \tfrac{1}{n}(U\Sigma V^\intercal)^\intercal U\Sigma V^\intercal \\ &= \tfrac{1}{n} V\Sigma^\intercal U^\intercal U\Sigma V^\intercal \\ &= \tfrac{1}{n} V\Sigma^\intercal I_m \Sigma V^\intercal \\ &= V\tfrac{1}{n}\Sigma^\intercal\Sigma V^\intercal \end{aligned}$$

Hence $Q = V$ and $\Lambda = \frac{1}{n}\Sigma^\intercal\Sigma = \frac{1}{n}\Sigma^2$. In particular, this means that the principal components are equal to the right singular vectors.

(d) With PCA, we project the original data X into a new space $Y = XV$ of dimension d. Compute the covariance matrix of Y.

With a slight abuse of notation treating V as square rather than $m \times d$:

$$\mathrm{Cov}[Y] = n^{-1}Y^TY = n^{-1}V^TX^TXV = n^{-1}V^TV\Lambda V^TV = n^{-1}\Lambda. \tag{1}$$

In other words, the covariance matrix for Y is diagonal, i.e., PCA decorrelates the principal components.