Exercise List: Dimension reduction and associated linear algebra

Robert M. Gower.

November 8, 2019

Notation: For every $x, y, \in \mathbb{R}^d$ let $\langle x, y\rangle \stackrel{\text{def}}{=} x^\top y$ and let $\|x\|_2 = \sqrt{\langle x, x\rangle}$.

1 Linear Algebra

Ex. 1 — Prove the follow lemma

Lemma 1.1. For any matrix $M \in \mathbb{R}^{d\times k}$ and $X \in \mathbb{R}^{d\times n}$ we have that

$$\mathbf{Null}\left(M^\top X\right) = \mathbf{Null}\left(MM^\top X\right) \tag{1}$$

and

$$\mathbf{Range}\left(X^\top MM^\top\right) = \mathbf{Range}\left(X^\top M\right). \tag{2}$$

Hint: Prove first (1) then consider the complement. Note that one of the inclusions in (1) is trivial, so you only need to prove the other inclusion.

Ex. 2 — Prove the following lemma

Lemma 1.2. For any matrix W and symmetric positive definite matrix G,

$$\mathbf{Null}\left(W\right) = \mathbf{Null}\left(W^\top GW\right) \tag{3}$$

and

$$\mathbf{Range}\left(W^\top\right) = \mathbf{Range}\left(W^\top GW\right). \tag{4}$$

Hint: Use that there exists positive definite $G^{1/2}$ such that $G^{1/2}G^{1/2} = G$.

2 Dimension Reduction

Ex. 3 — Let $X = [x_1, \ldots, x_n] \in \mathbb{R}^{d \times n}$ and suppose that it has been centered and scaled, that is consider the transformation given by $x_i \to \mathrm{diag}(\sigma_1, \ldots, \sigma_d)^{-1}(x_i - \mu)$ where

$$\mu_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij} \quad \text{and} \quad \sigma_i = \frac{1}{n}\sum_{i=1}^{n}(x_{ij} - \mu_j)^2, \quad \text{for } j = 1, \ldots, d.$$

Part I

Assume that $\sigma_1, \ldots, \sigma_n$ has been pre-calculated and saved. Find the matrices $\Sigma \in \mathbb{R}^{d \times d}$ and $D \in \mathbb{R}^{n \times n}$ such that the transformation $X \to \Sigma X D$ applies a centring and scaling to the data. Given a test data $X_t \in \mathbb{R}^{d \times k}$ can you apply this exact same transformation to X_t, that is $X_t \to \Sigma X_t D$?

Part II

Assume that $\sigma_1, \ldots, \sigma_n$ and $\mu_1, \ldots, \mu_n$ have been pre-calculated and saved. Find the matrix $\Sigma \in \mathbb{R}^{d \times d}$ and $C \in \mathbb{R}^{d \times n}$ such that $X \to \Sigma(X + C)$ applies a centring and scaling to the data. Given a test data $X_t \in \mathbb{R}^{d \times k}$ can you apply this exact same transformation to X_t?

Part III

Let P be the rank–r PCA transform of X. In other words assume that

$$P = \arg \max_{V \in \mathbb{R}^{d \times r},\, V^\top V = I} \|V^\top X\|_F^2. \tag{5}$$

Show that if X is centred and scaled then P also maximizes the scattering of the project data, that is

$$P \in \arg \max_{V \in \mathbb{R}^{d \times r},\, V^\top V = I} \sum_{j=1}^{n}\sum_{i=1}^{n} \|V^\top x_i - V^\top x_j\|_2^2 \tag{6}$$

Hint: For every $x, y \in \mathbb{R}^d$ and $V \in \mathbb{R}^{d \times r}$ we have that

$$\left\langle V^\top x, V^\top y \right\rangle_2 = \left\langle VV^\top, yx^\top \right\rangle = \mathrm{Tr}\left(VV^\top, yx^\top\right).$$

3 Sparse Matrix Formats

Write the pseudo code of an algorithm that:

Ex. 4 — Takes a matrix $M \in \mathbb{R}^{r \times d}$ in CSR format, a matrix $X \in \mathbb{R}^{d \times n}$ is the standard dense format, and returns $MX \in \mathbb{R}^{r \times n}$ in the standard dense format.

Ex. 5 — Takes a matrix $M \in \mathbb{R}^{n \times r}$ in CSR format, a matrix $X \in \mathbb{R}^{d \times n}$ is the standard dense format, and returns $XM \in \mathbb{R}^{d \times r}$ in the standard dense format.