Statistics 312/612, fall 2016
Homework # 3
Due: Monday 26 September

[1] (10 points) Suppose X is an $n \times p$ matrix of constants with rank p. Suppose also that $y = X\beta + \xi$ for an unknown $\beta \in \mathbb{R}^p$ and random ξ for which $\mathbb{E}\xi = 0$ and $\text{var}(\xi) = \sigma^2 I_n$. Let $\widehat{b}$ be the least squares estimator for β, as described on the handout MeanCov.pdf. Find

$$\frac{\max\{\text{var}(q^T\widehat{b}) : \|q\| = 1\}}{\min\{\text{var}(q^T\widehat{b}) : \|q\| = 1\}}.$$

[2] Let Z be an $m \times 1$ random vector with $\mathbb{E}Z = \mu_z$ and Y be an $n \times 1$ random vector with $\mathbb{E}Y = \mu_y$. Suppose both $V_z = \text{var}(Z)$ and $V_y = \text{var}(Y)$ are nonsingular. Find the linear functions a^TZ and b^TY for which the correlation is largest. That is, find vectors of constants $a \in \mathbb{R}^m$ and $b \in \mathbb{R}^n$ to maximize

$$\frac{\text{cov}(a^TZ, b^TY)}{\sqrt{\text{var}(a^TZ)\text{var}(b^TY)}}.$$

Follow these steps.

(i) (5 points) Explain why we may assume that both μ_z and μ_y are zero, so the problem becomes:

$$\text{find } a \text{ and } b \text{ to maximize } \frac{a^T\mathbb{E}(ZY^T)b}{\sqrt{\mathbb{E}(a^TZ)^2\ \mathbb{E}(b^TY)^2}}$$

(ii) (5 points) Explain why the problem is essentially unchanged if we replace Z by $\widetilde{Z} = V_z^{-1/2}Z$ and Y by $\widetilde{Y} = V_y^{-1/2}Y$. See the back of this problem sheet if you are not familiar with the positive square root of a symmetric, positive definite square matrix.

(iii) (5 points) Write K for $\mathbb{E}(\widetilde{Z}\widetilde{Y}^T)$. Show that the problem reduces to a search for unit vectors u and v for which u^TKv is maximized.

(iv) (5 points) Use the singular value decomposition for K to complete the solution.

[3] Suppose vectors in $\mathbb{R}^6$ are arranged as 2×3 tables: $y = \boxed{\begin{matrix} y_{1,1} & y_{1,2} & y_{1,3} \\ y_{2,1} & y_{2,2} & y_{2,3} \end{matrix}}$ and

$$\mathbb{1} = \boxed{\begin{matrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{matrix}} \qquad R_1 = \boxed{\begin{matrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{matrix}} \qquad R_2 = \boxed{\begin{matrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{matrix}}$$

$$C_1 = \boxed{\begin{matrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{matrix}} \qquad C_2 = \boxed{\begin{matrix} 0 & 1 & 0 \\ 0 & 1 & 0 \end{matrix}} \qquad C_3 = \boxed{\begin{matrix} 0 & 0 & 1 \\ 0 & 0 & 1 \end{matrix}}$$

Let $\mathcal{X}$ denote the subspace of $\mathbb{R}^6$ spanned by $\{\mathbb{1}, R_1, R_2, C_1, C_2, C_3\}$ and

$$\mathcal{X}_1 = \text{span}\{\mathbb{1}\} \qquad \mathcal{X}_R = \text{span}\{R_1 - \tfrac{1}{2}\mathbb{1}, R_2 - \tfrac{1}{2}\mathbb{1}\}$$
$$\mathcal{X}_C = \text{span}\{C_1 - \tfrac{1}{3}\mathbb{1}, C_2 - \tfrac{1}{3}\mathbb{1}, C_3 - \tfrac{1}{3}\mathbb{1}\}.$$

(i) (5 points) Show that the subspaces $\mathcal{X}_{\mathbb{1}}, \mathcal{X}_R, \mathcal{X}_C$ are mutually orthogonal with dimensions $1, 1, 2$.

(ii) (5 points) Show that the orthogonal projection of a vector y in $\mathbb{R}^6$ onto $\mathcal{X}_C$ equals $\sum_{j=1}^3 \bar{y}_{\cdot,j}\left(C_j - \frac{1}{3}\mathbb{1}\right)$ where $\bar{y}_{\cdot,j} = \frac{1}{2}(y_{1,j} + y_{2,j})$. Hint: Write out the 2×3 tables.

(iii) (10 points) Show that the orthogonal projection of y onto $\mathcal{X}$ equals

$$\bar{y}_{\cdot\cdot}\mathbb{1} + \sum\nolimits_{j=1}^3 \bar{y}_{\cdot,j}\left(C_j - \tfrac{1}{3}\mathbb{1}\right) + \sum\nolimits_{i=1}^2 \bar{y}_{i\cdot}(R_i - \tfrac{1}{2}\mathbb{1}),$$

where a dot in the subscript indicates an averaging over the missing subscript.

Spectral decomposition (Axler, 2015, Section 7.B)

Suppose S is a $k \times k$ matrix of real numbers. An ***eigenvector*** of S is a nonzero vector x in $\mathbb{R}^k$ for which $Sx = \theta x$ for some constant θ. The value θ is called an eigenvalue.

If S is symmetric (that is, $S^T = S$) then there exists an orthonormal basis $q_1, \dots, q_k$ for $\mathbb{R}^k$ such that each q_i is an eigenvector with the corresponding eigenvalues all real. That is, the matrix $Q = (q_1, \dots, q_k)$ is orthogonal and

$$SQ = Q\Theta \qquad \text{where } \Theta = \operatorname{diag}(\theta_1, \dots, \theta_k).$$

Equivalently,

$$S = Q\Theta Q^T = \sum\nolimits_{i \le k} \theta_i q_i q_i^T,$$

which is just like a singular-value decomposition with $U = V$.

> **Remark.** In fact, the svd of a matrix X can be derived from the spectral decomposition of the symmetric matrix $X^T X$.

The matrix S is said to be non-negative definite if $x^T S x \ge 0$ for every x in $\mathbb{R}^k$. For such a matrix $0 \le q_i^T S q_i = \theta_i$ for every i. The matrix

$$T = \sum\nolimits_{i \le k} \sqrt{\theta_i} q_i q_i^T = Q\Theta^{1/2} Q^T$$

is also symmetric and positive definite, with $T^2 = S$. Sometimes T is called a ***positive square root*** of S (Axler, 2015, 7.35). Some authors write $S^{1/2}$ for T.

The matrix S is nonsingular if and only if all the θ_i are > 0. In that case, T is also nonsingular, with inverse

$$T^{-1} = Q\Theta^{-1/2} Q^T = \sum\nolimits_{i \le k} \theta^{-1/2} q_i q_i^T.$$

Some authors write $S^{-1/2}$ for T^{-1}. It has the property that

$$T^{-1} S T^{-1} = Q\Theta^{-1/2} Q^T Q\Theta Q^T Q\Theta^{-1/2} Q^T = I_k.$$

<1> **Example.** Suppose Z is a $k \times 1$ random vector with $V = \operatorname{var}(Z)$ well defined. For simplicity, suppose $\mathbb{E}Z = 0$ so that $V = \mathbb{E}(ZZ^T)$. Clearly V is symmetric. It is also non-negative definite because $x^T V x = \mathbb{E}\left((x^T Z)^2\right) \ge 0$. It has a positive square root T. If V is nonsingular then there exists a symmetric matrix $V^{-1/2}$ for which $V^{-1/2} V V^{-1/2} = I_k$. Consequently, $\operatorname{var}(V^{-1/2} Z) = I_k$.

□

References

Axler, S. J. (2015). *Linear Algebra Done Right* (Third ed.). Undergraduate Texts in Mathematics. Springer.