Problem Set 3: Matrix Calculus

Matrices are typically used to represent big systems of linear equations. Of'course one can use calculus in such systems. Let $f(\mathbf{X})$ be a scalar function of the $(n \times m)$ matrix $\mathbf{X}$. Then $\partial f(\mathbf{X}) / \partial \mathbf{X}$ is an $(n \times m)$ matrix whose (i, j)-th element is $\partial f(\mathbf{X}) / \partial x_{ij}$.

Exercise 1 *Let* $\mathbf{x}, \mathbf{a}$ *denote two* $(n \times 1)$ *vectors. Denote* $f(\mathbf{x}) = \mathbf{a}'\mathbf{x}$. *Show that*

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \mathbf{a}$$

Exercise 2 *If* $\mathbf{A}$ *denotes* $(n \times m)$ *matrix and* $\mathbf{x}$ *is* $(m \times 1)$ *vector, verify that*

$$\frac{\partial (\mathbf{A}\mathbf{x})}{\partial \mathbf{x}'} = \mathbf{A}$$

Exercise 3 *Take* $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 3 & 1 \end{bmatrix}$. *Verify*

$$\frac{\partial \mathbf{x}'\mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = (\mathbf{A} + \mathbf{A}')\,\mathbf{x}$$

where $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ *is an* (2×1) *vector.*

Exercise 4 *Take matrix*

$$\mathbf{X} = \begin{bmatrix} 2 & 3 \\ 4 & 5 \\ -1 & 0 \end{bmatrix}$$

Compute $\mathbf{X}'\mathbf{X}$.

Exercise 5 *Show that* $\mathbf{X}'\mathbf{X}$ *is symmetric. Compute*

$$\frac{\partial \mathbf{b}' (\mathbf{X}'\mathbf{X})\, \mathbf{b}}{\partial \mathbf{b}}$$

using formula from excercise 3.

Exercise 6 *Suppose that for given* $(n \times 1)$ *vector* $\mathbf{y}$ *and* $(n \times k)$ *full column rank matrix* $\mathbf{X}$ *we want to find* $(k \times 1)$ *vector of coefficients* $\mathbf{b}$, *such that the linear combination of column vectors of* $\mathbf{X}$ *(i.e.* $\mathbf{X}\mathbf{b}$*) is closest to vector* $\mathbf{y}$, *i.e. we want to minimise* $\|\mathbf{y} - \mathbf{X}\mathbf{b}\|^2 \equiv \langle \mathbf{y} - \mathbf{X}\mathbf{b}, \mathbf{y} - \mathbf{X}\mathbf{b} \rangle = (\mathbf{y} - \mathbf{X}\mathbf{b})' (\mathbf{y} - \mathbf{X}\mathbf{b})$. *We can proceed in two ways:*

(a) *Use calculus: Compute* $\frac{\partial \|\mathbf{y} - \mathbf{X}\mathbf{b}\|^2}{\partial \mathbf{b}'}$ *and find* $\widehat{\mathbf{b}}$ *for which this derivation is zero.*

(b) *Use so called orthogonal projection theorem, which says that:*

$$\widehat{\mathbf{b}} = \arg\min_{\mathbf{b}\in\mathbb{R}^k} \|\mathbf{y} - \mathbf{X}\mathbf{b}\|^2 \iff \left(\mathbf{y} - \mathbf{X}\widehat{\mathbf{b}}\right) \text{ is orthogonal to any linear combination } \mathbf{X}\mathbf{b},\ \mathbf{b} \in \mathbb{R}^k$$

I.e. we have to find vector $\widehat{\mathbf{b}}$*, which satisfies following orthogonality conditions*

$$\left\langle \mathbf{y} - \mathbf{X}\widehat{\mathbf{b}}, \mathbf{X} \right\rangle \equiv \left(\mathbf{y} - \mathbf{X}\widehat{\mathbf{b}}\right)' \mathbf{X} = \mathbf{0}_k$$