Exercise 6.6.14

Linear Algebra
MATH 2130

SEBASTIAN CASALAINA

ABSTRACT. This is Exercise 6.6.14 from Lay [LLM16, §6.6]:

Exercise 6.6.14. Let $\overline{x} = \frac{1}{m}(x_1 + \cdots + x_m)$ and let $\overline{y} = \frac{1}{n}(y_1 + \cdots + y_m)$. *Show that the least-squares line for the data* $(x_1, y_1), \ldots, (x_m, y_m)$ *must pass through* $(\overline{x}, \overline{y})$. That is, show that $\overline{x}$ and $\overline{y}$ must satisfy the linear equation $\overline{y} = \hat{\beta}_0 + \hat{\beta}_1 \overline{x}$.

[*Hint*: Derive this equation from the vector equation $\mathbf{y} = X\hat{\boldsymbol{\beta}} + \boldsymbol{\epsilon}$. Denote the first column of X by $\mathbf{1}$. Use the fact that the residual vector $\boldsymbol{\epsilon}$ is orthogonal to the column space of X and hence is orthogonal to $\mathbf{1}$.]

Solution. Recall that given an $m \times n$ matrix A and a vector $\mathbf{b}$ in $\mathbb{R}^m$, a least squares solution $\hat{\mathbf{x}}$ to the equation

$$A\mathbf{x} = \mathbf{b}$$

is a vector $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that $A\hat{\mathbf{x}}$ is equal to $\pi_W(\mathbf{b})$, where $W = \mathrm{Im}(A) = \mathrm{Col}(A) \subseteq \mathbb{R}^m$ is the image of the linear map associated to A, and π_W is the orthogonal projection of $\mathbb{R}^m$ onto W. From the definition of the orthogonal projection, we have that $\mathbf{b} - \pi_W(\mathbf{b})$ is in $W^\perp$ $(= (\mathrm{Col}(A))^\perp)$. If we define the "error", or "residual vector", $\boldsymbol{\epsilon}$ as

$$\boldsymbol{\epsilon} := \mathbf{b} - A\hat{\mathbf{x}},$$

then since $A\hat{\mathbf{x}} = \pi_W(\mathbf{b})$, we have $\boldsymbol{\epsilon} = \mathbf{b} - A\hat{\mathbf{x}} = \mathbf{b} - \pi_W(\mathbf{b}) \in W^\perp = (\mathrm{Col}(A))^\perp$.

Date: November 12, 2022.

Now, translating this into the language of least-squares line fitting, we have

$$A = X = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix}$$

$$\mathbf{b} = \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

$$\hat{\mathbf{x}} = \hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix}$$

In this notation, the general assertion we explained above, that $\boldsymbol{\epsilon} = \mathbf{b} - A\hat{\mathbf{x}} = \mathbf{b} - \pi_W(\mathbf{b}) \in W^\perp = (\mathrm{Col}(A))^\perp$, means that $\boldsymbol{\epsilon} = \mathbf{y} - X\hat{\boldsymbol{\beta}} \in (\mathrm{Col}(X))^\perp$; this was the first assertion the hint suggested we prove. In particular, $\boldsymbol{\epsilon}$ is orthogonal to the first column of X, which is the vector denoted by $\mathbf{1}$.

In other words, we have $\langle \mathbf{1}, \boldsymbol{\epsilon} \rangle = 0$, i.e., we have

$$\begin{aligned} 0 &= \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \left(\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} - \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} \right) \\ &= \sum_{i=1}^m y_i - \sum_{i=1}^m (\hat{\beta}_0 + \hat{\beta}_1 x_i) \end{aligned}$$

Dividing by m, we have that

$$0 = \overline{y} - (\hat{\beta}^0 + \hat{\beta}_1 \overline{x}),$$

which was what we were trying to prove. $\square$

References

[LLM16] David Lay, Stephen Lay, and Judi McDonald, *Linear Algebra and its Applications*, Fifth edition, Pearson, 2016.

University of Colorado, Department of Mathematics, Campus Box 395, Boulder, CO 80309

Email address: casa@math.colorado.edu