Statistics 244 — Fall 2025 — Assignment 1

Due Friday, September 19, 2025

Homework is to be uploaded to Gradescope by 10:00pm on Friday evening.

Readings:

"Introductory lecture" and "Linear algebra issues in linear models" course notes.

Agresti: Chapter 1

Written assignment

1. Suppose $\boldsymbol{X}$ is an $n \times p$ model matrix. Show that $\{\boldsymbol{X\beta}|\boldsymbol{\beta} \in \mathbb{R}^p\}$ is a vector space.

   Solution: Let $V = \{\boldsymbol{X\beta}|\boldsymbol{\beta} \in \mathbb{R}^p\}$. We simply need to show that that set of vectors is closed under linear combination, i.e a) $\boldsymbol{u} + \boldsymbol{v} \in V$ and $c\boldsymbol{u} \in V$. Since we are looking for $\beta \in \mathbb{R}^p$, we are simply talking about any arbitrary linear combination of the columns of $\boldsymbol{X}$. Let us first prove the first property:

   (a) if $\boldsymbol{u} \in V$ and $\boldsymbol{v} \in V$, $\boldsymbol{u} + \boldsymbol{v} \in V$. By the definition of V, $\exists \beta_1, \beta_2$ such that $\boldsymbol{X}\beta_1 = \boldsymbol{u}$ and $\boldsymbol{X}\beta_2 = \boldsymbol{v}$. In that case, we can rewrite their sum as $\boldsymbol{X}(\beta_1 + \beta_2)$. By the definition of β, as $\beta_1 + \beta_2 \in \mathbb{R}^p$, $\boldsymbol{u} + \boldsymbol{v} \in V$.

   (b) if $\boldsymbol{u} \in V$, then $c\boldsymbol{u} \in V$. This is even simpler to show. By definition of V, $\exists \beta_3$ such that $\boldsymbol{X}\beta_3 = \boldsymbol{u}$. Then, $c\boldsymbol{u} = c\boldsymbol{X}\beta_3 = \boldsymbol{X}c\beta_3$, and as c is a real scalar, $c\beta_3 \in \mathbb{R}^p$.

   Thus, $\{\boldsymbol{X\beta}|\boldsymbol{\beta} \in \mathbb{R}^p\}$ is a vector space.

2. (a) For $n \times p$ model matrix $\boldsymbol{X}$, show that the null space of $\boldsymbol{X}$, $N(\boldsymbol{X})$, is the orthogonal complement of the column space of $\boldsymbol{X}^\top$, that is, $C(\boldsymbol{X}^\top)^\perp$.

Solution: The definition of orthogonal complement is

$$\forall \boldsymbol{u} \in W, \forall \boldsymbol{v} \in W^{\perp} \subseteq \mathbb{R}^n \iff \boldsymbol{u}^\top \boldsymbol{v} = 0$$

Here, if the equation on the right is satisfied, then $W^\perp$ is the orthogonal complement of W, and vice versa.
As for null space, it is defined as

$$N(\boldsymbol{X}) = \{\boldsymbol{\zeta} : \boldsymbol{X}\boldsymbol{\zeta} = \boldsymbol{0}\}$$

This means that the null space of $\boldsymbol{X}$ is simply the set of vectors for which $\boldsymbol{\zeta}$ is orthogonal to all rows in $\boldsymbol{X}$, as their multiplication should produce the $\boldsymbol{0}$, i.e the zero vector.
Let $\boldsymbol{k}^\top$ and $\boldsymbol{\ell}^\top$ be two rows from $\boldsymbol{X}$. We want to show being orthogonal to each separately, i.e $\boldsymbol{k}^\top\boldsymbol{\zeta} = \boldsymbol{0}$ and $\boldsymbol{\ell}^\top\boldsymbol{\zeta} = \boldsymbol{0}$, means being orthogonal to any arbitrary linear combination too. This follows trivially:

$$(c\boldsymbol{k}^\top + p\boldsymbol{\ell}^\top)\boldsymbol{\zeta} = c\boldsymbol{k}^\top\boldsymbol{\zeta} + p\boldsymbol{\ell}^\top\boldsymbol{\zeta} = c \cdot 0 + p \cdot 0 = 0$$

Then, it can be said that $N(\boldsymbol{X})$ is orthogonal to all linear combinations of the rows of $\boldsymbol{X}$. As the column space of $\boldsymbol{X}^\top$ is simply the vector space created by the span of the columns of $\boldsymbol{X}^\top$, and the span of the columns of $\boldsymbol{X}^\top$ is equivalent to the span of the rows of $\boldsymbol{X}$, $N(\boldsymbol{X}) = C(X^\top)^\perp$.

(b) Let $\boldsymbol{X} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}$, and let $V = C(\boldsymbol{X})$ be the vector space spanned by the columns of $\boldsymbol{X}$. Determine the orthogonal complement $V^\perp$ of V as a closed-form expression. Use the result of part (a) as a guide.

Solution: We saw in part (a) that the null space of $\boldsymbol{X}$ is equivalent to the orthogonal complement of $C(X^\top)$. Here, we want to look at it from the other way. The orthogonal complement $V^\perp$ of V can be written as the null space of the transpose of the model matrix that creates it, i.e

$$V^\perp = N(\boldsymbol{X}^\top)$$

Considering the simplicity of $\boldsymbol{X}$, this is fairly simple:

$$N(\boldsymbol{X}^\top) = \{\boldsymbol{\zeta} : \boldsymbol{X}^\top\boldsymbol{\zeta} = \boldsymbol{0}\} = \{\boldsymbol{\zeta} : \boldsymbol{\zeta} = c(-1, 1, -1)^\top \forall c \in \mathbb{R}\} = \text{span}((-1, 1, -1)^\top)$$

3. A model M has model matrix $\boldsymbol{X}$. A simpler model M_0 results from removing the final term in M, and hence has model matrix $\boldsymbol{X}_0$ that deletes the final column from $\boldsymbol{X}$. From the definition of a column space, explain why $C(\boldsymbol{X}_0) \subseteq C(\boldsymbol{X})$.

Solution: Looking back at the definition of column space, we see

$$C(\boldsymbol{X}) = \{\boldsymbol{\eta} : \boldsymbol{\eta} = \boldsymbol{X}\beta, \forall \beta \in \mathbb{R}^p\} \subseteq \mathbb{R}^n$$

We want to show that $C(\boldsymbol{X}_0)$ is either a subset of or equal to $C(\boldsymbol{X})$. An important detail here to notice is that the model matrix $\boldsymbol{X}$ isn't said to be full-rank, which is where the possibility of $\boldsymbol{X}_0$ and $\boldsymbol{X}$ having the same column space comes from. Here, we have two options: a) Either a linearly independent vector was removed, in which case $\boldsymbol{X}_0$ is a subset, or b) a dependent vector was removed, in which case $\boldsymbol{X}_0$ is the same as $\boldsymbol{X}$. If we prove these two statements, then we will have proven the statement in the question. However, another way to approach this which is agnostic of whether $\boldsymbol{x}_p$ was linearly independent or not is showing that if something is a member of $C(\boldsymbol{X}_0)$, then it also has to be a member of $C(\boldsymbol{X})$. First, we can rewrite the definition of the column space as such:

$$C(\boldsymbol{X}) = \{\boldsymbol{\eta} : \boldsymbol{\eta} = \boldsymbol{x}_1\beta_1 + \boldsymbol{x}_2\beta_2 + \cdots + \boldsymbol{x}_p\beta_p, \forall \beta_i \in \mathbb{R}\} \subseteq \mathbb{R}^n$$

Then, the column space of $\boldsymbol{X}_0$ is defined as:

$$C(\boldsymbol{X}_0) = \{\boldsymbol{\eta} : \boldsymbol{\eta} = \boldsymbol{x}_1\beta_1 + \boldsymbol{x}_2\beta_2 + \cdots + \boldsymbol{x}_{p-1}\beta_{p-1}, \forall \beta_i \in \mathbb{R}\} \subseteq \mathbb{R}^n$$

It is fairly evident from this example that

$$C(\boldsymbol{X}_0) \subseteq C(\boldsymbol{X})$$

as when β_p is zero, the set of vectors that can be expressed by both become equal, but $C(\boldsymbol{X})$ is strictly larger when β_p isn't zero.

4. Suppose that $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are full-rank $n \times p$ model matrices with $p < n$.
   (a) Show that if $\boldsymbol{A} \in \mathbb{R}^{p \times p}$ is nonsingular, $C(\boldsymbol{X}_1) = C(\boldsymbol{X}_1\boldsymbol{A})$.
   (b) Now suppose that $C(\boldsymbol{X}_1) = C(\boldsymbol{X}_2)$. This part of the problem involves showing that this implies the existence of a nonsingular matrix $\boldsymbol{A} \in \mathbb{R}^{p \times p}$ such that $\boldsymbol{X}_1 = \boldsymbol{X}_2\boldsymbol{A}$.
      i. To establish the existence of $\boldsymbol{A}$, let

$$\boldsymbol{X}_1 = [\boldsymbol{x}_{11} \mid \boldsymbol{x}_{12} \mid \cdots \mid \boldsymbol{x}_{1p}], \qquad \boldsymbol{A} = [\boldsymbol{a}_1 \mid \cdots \mid \boldsymbol{a}_p],$$

      where $\boldsymbol{x}_{1j}$ is the j-th column of $\boldsymbol{X}_1$, and $\boldsymbol{a}_j \in \mathbb{R}^p$ is the j-th column of $\boldsymbol{A}$. Show that each $\boldsymbol{x}_{1j}$ can be written as $\boldsymbol{x}_{1j} = \boldsymbol{X}_2\boldsymbol{a}_j$, and hence $\boldsymbol{X}_1 = \boldsymbol{X}_2\boldsymbol{A}$.
      ii. With the result from part (i), show that $\boldsymbol{A}$ must be invertible. Prove this by contradiction: assume that $\boldsymbol{A}$ is singular, so there exists a nonzero vector $\boldsymbol{v} \in \mathbb{R}^p$ such that $\boldsymbol{A}\boldsymbol{v} = \boldsymbol{0}$. Proceed by showing that this contradicts the assumption that $\boldsymbol{X}_1$ has full column rank.

5. Consider the model for the *two-way layout* for categorical variables A and B:

$$\mathrm{E}(y_{ijk}) = \beta_0 + \beta_i + \gamma_j, \quad \text{where } i = 1, \ldots, r; \quad j = 1, \ldots, c; \quad \text{and } k = 1, \ldots, n.$$

This model is *balanced*, having an equal sample size n in each of the rc cells. We might be interested in estimating the parameter vector $(\beta_0, \beta_1, \ldots, \beta_r, \gamma_1, \ldots, \gamma_c)$.

(a) For the model as stated, is the parameter vector identifiable? Why or why not? It might help to code the model for $\mathrm{E}(y_{ijk})$ into a model matrix that multiplies the parameter vector stated above.

(b) Give an example of a quantity that is a function of the model parameters, $\beta_0, \beta_1, \ldots, \beta_r, \gamma_1, \ldots, \gamma_c$, that is (i) not estimable, and (ii) estimable. In each case, provide a rigorous justification.

6. Suppose $\boldsymbol{y} \in \mathbb{R}^n$, and $\mathrm{E}(\boldsymbol{y}) = \boldsymbol{X}\boldsymbol{\beta}$ for $n \times p$ model matrix $\boldsymbol{X}$ and $\boldsymbol{\beta} \in \mathbb{R}^p$. This question concerns proving that $\boldsymbol{\ell}^\top \boldsymbol{\beta}$ is estimable if and only if

$$\boldsymbol{\ell}^\top = \boldsymbol{\ell}^\top (\boldsymbol{X}^\top \boldsymbol{X})^- (\boldsymbol{X}^\top \boldsymbol{X})$$

for selected $\boldsymbol{\ell} \in \mathbb{R}^p$.

(a) Prove that if $\boldsymbol{\ell}^\top \boldsymbol{\beta}$ is estimable, then for any g-inverse $(\boldsymbol{X}^\top \boldsymbol{X})^-$ of $\boldsymbol{X}^\top \boldsymbol{X}$,

$$\boldsymbol{\ell}^\top = \boldsymbol{\ell}^\top (\boldsymbol{X}^\top \boldsymbol{X})^- (\boldsymbol{X}^\top \boldsymbol{X})$$

for selected $\boldsymbol{\ell} \in \mathbb{R}^p$.

(b) Prove that if

$$\boldsymbol{\ell}^\top = \boldsymbol{\ell}^\top (\boldsymbol{X}^\top \boldsymbol{X})^- (\boldsymbol{X}^\top \boldsymbol{X})$$

for selected $\boldsymbol{\ell} \in \mathbb{R}^p$ and for some g-inverse $(\boldsymbol{X}^\top \boldsymbol{X})^-$ of $\boldsymbol{X}^\top \boldsymbol{X}$, then $\boldsymbol{\ell}^\top \boldsymbol{\beta}$ is estimable.

(c) For full-rank $\boldsymbol{X}$, what do the previous results imply about the choice of $\boldsymbol{\ell} \in \mathbb{R}^p$ for $\boldsymbol{\ell}^\top \boldsymbol{\beta}$ to be estimable?

7. Suppose $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ are two arbitrary subspaces of $\mathbb{R}^n$. The goal of this problem is to prove

$$(\boldsymbol{W}_1 \cap \boldsymbol{W}_2)^\perp = \boldsymbol{W}_1^\perp + \boldsymbol{W}_2^\perp,$$

where, in general, for $V_1, V_2 \subseteq \mathbb{R}^n$,

$$V_1 + V_2 = \{v_1 + v_2 : v_1 \in V_1, v_2 \in V_2\}.$$

You may assume below (without proof) that for any subspace $A \subseteq \mathbb{R}^n$, $(A^\perp)^\perp = A$.

(a) We first need to prove three lemmas.

i. Prove that, for any subspace $A \subseteq \mathbb{R}^n$, $(A^\perp)^\perp = A$.

ii. Show that, for any $A \subseteq \mathbb{R}^n$, $\mathbf{0} \in A^\perp$.

iii. Let $A, B \subseteq \mathbb{R}^n$. Prove that $A \subseteq B$ if and only if $B^\perp \subseteq A^\perp$.

(b) To prove the main result of the problem, we first want to show that

$$(\boldsymbol{W}_1 \cap \boldsymbol{W}_2)^\perp \subseteq \boldsymbol{W}_1^\perp + \boldsymbol{W}_2^\perp.$$

i. Show that $\boldsymbol{W}_i^\perp \subseteq \boldsymbol{W}_1^\perp + \boldsymbol{W}_2^\perp$ for $i = 1, 2$.

ii. Show that the above implies

$$(\boldsymbol{W}_1^\perp + \boldsymbol{W}_2^\perp)^\perp \subseteq (\boldsymbol{W}_1 \cap \boldsymbol{W}_2),$$

and therefore $(\boldsymbol{W}_1 \cap \boldsymbol{W}_2)^\perp \subseteq \boldsymbol{W}_1^\perp + \boldsymbol{W}_2^\perp$.

(c) Now show the converse of part (b), namely that

$$\boldsymbol{W}_1^\perp + \boldsymbol{W}_2^\perp \subseteq (\boldsymbol{W}_1 \cap \boldsymbol{W}_2)^\perp.$$

Hint: Choose $v_1 \in \boldsymbol{W}_1^\perp$ and $v_2 \in \boldsymbol{W}_2^\perp$, and proceed from there.