Statistical Sciences Honours
Matrix Methods
Lecture 3 – Inverses

Stefan S. Britz
stefan.britz@uct.ac.za

Department of Statistical Sciences
University of Cape Town

3.1 Notation

- In the previous lectures we saw that matrix multiplication is just a set of linear transformations.
- Now we ask whether these transformations can be reversed, or rather inverted.
- For any square matrix $\boldsymbol{A}$, suppose that there exists another matrix $\boldsymbol{A}^{-1}$ of the same size such that $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}$.
- We term $\boldsymbol{A}^{-1}$ the inverse of $\boldsymbol{A}$.

3.1 Notation

- For any set of linear equations, if the inverse exists, then

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b} \\ \boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{x} &= \boldsymbol{A}^{-1}\boldsymbol{b} \\ \boldsymbol{x} &= \boldsymbol{A}^{-1}\boldsymbol{b} \end{aligned}$$

- Therefore, solving linear equations and inverting square matrices are equivalent operations.
- We will look at an application example shortly, first we need to know how to determine the inverse.

3.2 Evaluation using elementary operators

- To get a feel, first try the following exercise:

Exercise 3.1

Show that $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$

3.2 Evaluation using elementary operators

- To get a feel, first try the following exercise:

Exercise 3.1

Show that $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$

- Now suppose we can find a sequence of elementary operator matrices $\boldsymbol{E}_{(1)}, \boldsymbol{E}_{(2)}, \ldots, \boldsymbol{E}_{(K)}$ such that

$$\boldsymbol{E}_{(K)}\boldsymbol{E}_{(K-1)}\cdots\boldsymbol{E}_{(2)}\boldsymbol{E}_{(1)}\boldsymbol{A} = \boldsymbol{I}$$

- In other words if we apply a sequence of elementary row operations to $\boldsymbol{A}$, then the result is the identity matrix.

3.2 Evaluation using elementary operators

- This implies by definition that $\boldsymbol{E}_{(K)}\boldsymbol{E}_{(K-1)}\cdots\boldsymbol{E}_{(2)}\boldsymbol{E}_{(1)} = \boldsymbol{A}^{-1}$.
- We do not need to explicitly know the matrices $\boldsymbol{E}_{(\boldsymbol{k})}$, since by definition

$$\boldsymbol{E}_{(K)}\boldsymbol{E}_{(K-1)}\cdots\boldsymbol{E}_{(2)}\boldsymbol{E}_{(1)}\boldsymbol{I} = \boldsymbol{A}^{-1}$$

- Therefore, if we apply the same order of row operations to the identity matrix, this will yield $\boldsymbol{A}^{-1}$.
- This process, known as Gaussian Elimination, or row reduction, is tedious to do by hand. Exercise 3.2 is optional if you would like to try!
- Note that the same reasoning holds for column operations, which would result from post-multiplication of elementary operators.

3.3 Adjugate matrix

- Another way of determining the inverse (although not much less tedious) is to use the adjugate matrix.
- The adjugate of $\boldsymbol{A}$, or $adj(\boldsymbol{A})$, is the transpose of the matrix in which each element of $\boldsymbol{A}$ is replaced by its cofactor.
- We then determine the inverse as

$$\boldsymbol{A}^{-1} = \frac{adj(\boldsymbol{A})}{|\boldsymbol{A}|}$$

- This implies an important property of matrices: The inverse of the matrix $\boldsymbol{A}$ exists if and only if $|\boldsymbol{A}| \neq 0$.
- If $|\boldsymbol{A}| = 0$, then we say that the matrix is **singular**.

3.3 Adjugate matrix

This approach allows us to easily determine the inverse of a 2×2 matrix.

Let $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

3.3 Adjugate matrix

This approach allows us to easily determine the inverse of a 2×2 matrix.

Let $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

We know that $|\boldsymbol{A}| = ad - bc$

By inspection we can see that the adjugate is $adj(\boldsymbol{A}) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

Therefore, we have

$$\boldsymbol{A}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Class exercise

Let's use this knowledge to solve a simple problem.

Suppose there are two boys, Xola and Xolisa, whose ages are unknown. We do know that Xola is two years younger than Xolisa, but if you double (only) his age, he would be 3 years older than Xolisa. Determine their ages.

Class exercise

Exercise 3.4

Check the following relations

(a) $[a\boldsymbol{I}+b\boldsymbol{J}]^{-1}=\frac{1}{a}\left[\boldsymbol{I}-\frac{b}{a+nb}\boldsymbol{J}\right]$ where n is the size of $\boldsymbol{I}$ and $\boldsymbol{J}$, and where $a\neq 0$ and $a+nb\neq 0$.

(b) $\left(\boldsymbol{I}+\boldsymbol{A}^{-1}\right)^{-1}=\left[\left(\boldsymbol{A}+\boldsymbol{I}\right)\boldsymbol{A}^{-1}\right]^{-1}=\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{I})^{-1}$

(c) $(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{B}')^{-1}\boldsymbol{B}=\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{I}+\boldsymbol{B}'\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}$

(d) $(\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1})^{-1}=\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{B}=\boldsymbol{B}(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{A}$

(e) $\boldsymbol{A}-\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{A}=\boldsymbol{B}-\boldsymbol{B}(\boldsymbol{A}+\boldsymbol{B})^{-1}\boldsymbol{B}$

(f) $\boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}=\boldsymbol{A}^{-1}(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{B}^{-1}$

(g) $(\boldsymbol{I}+\boldsymbol{A}\boldsymbol{B})^{-1}=\boldsymbol{I}-\boldsymbol{A}(\boldsymbol{I}+\boldsymbol{B}\boldsymbol{A})^{-1}$

(h) $(\boldsymbol{I}+\boldsymbol{A}\boldsymbol{B})^{-1}\boldsymbol{A}=\boldsymbol{A}(\boldsymbol{I}+\boldsymbol{B}\boldsymbol{A})^{-1}$

Exercises

- The relationships themselves are not important, the goal is to practise manipulating equations containing inverses.
- You are only required to **check** the validity of each statement, not prove it. So just show that the LHS = RHS. Example a):

  If

  $$[a\boldsymbol{I} + b\boldsymbol{J}]^{-1} = \frac{1}{a}\left[\boldsymbol{I} - \frac{b}{a+nb}\boldsymbol{J}\right]$$

  then

  $$[a\boldsymbol{I} + b\boldsymbol{J}]\,[a\boldsymbol{I} + b\boldsymbol{J}]^{-1} = \boldsymbol{I} = [a\boldsymbol{I} + b\boldsymbol{J}]\,\frac{1}{a}\left[\boldsymbol{I} - \frac{b}{a+nb}\boldsymbol{J}\right]$$

  Now you only need to show that $[a\boldsymbol{I} + b\boldsymbol{J}]\,\frac{1}{a}\left[\boldsymbol{I} - \frac{b}{a+nb}\boldsymbol{J}\right] = \boldsymbol{I}$

  Many of the other exercises can be approached in a similar way.

3.4 Completing the square (univariate)

- Completing the square (univariate) is a technique used to rewrite a quadratic expression $ax^2 + bx + c$ in a more convenient form.
- This form has the variable contained in a squared term and the appropriate constants added.
- The completed square form can be written as:

$$ax^2 + bx + c = a\left(x + \frac{b}{2a}\right)^2 + \left(c - \frac{b^2}{4a}\right)$$

- We can now extend this idea to the multivariate case.

3.4 Completing the square (multivariate)

Let $\boldsymbol{A}: p \times p$ be a non-singular symmetric matrix, $\boldsymbol{x}$ and $\boldsymbol{b}$ be p-dimensional vectors, and c a scalar.

To complete the square, we aim to express the multivariate expression $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} - 2\boldsymbol{b}'\boldsymbol{x} + c$ as a quadratic form plus a term of constants, as follows:

$$(\boldsymbol{x} - \boldsymbol{d})'\boldsymbol{A}(\boldsymbol{x} - \boldsymbol{d}) + e,$$

with appropriate vector $\boldsymbol{d}$ and constant e.

We will only state the result, without derivation:

$$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} - 2\boldsymbol{b}'\boldsymbol{x} + c = \left(\boldsymbol{x} - \boldsymbol{A}^{-1}\boldsymbol{b}\right)'\boldsymbol{A}\left(\boldsymbol{x} - \boldsymbol{A}^{-1}\boldsymbol{b}\right) - \boldsymbol{b}'\boldsymbol{A}^{-1}\boldsymbol{b} + c$$

3.4 Completing the square (multivariate)

This result can easily be verified by starting with the RHS and multiplying out the quadratic form:

$$
\begin{aligned}
& \left(\boldsymbol{x}-\boldsymbol{A}^{-1}\boldsymbol{b}\right)'\boldsymbol{A}\left(\boldsymbol{x}-\boldsymbol{A}^{-1}\boldsymbol{b}\right)-\boldsymbol{b}'\boldsymbol{A}^{-1}\boldsymbol{b}+c \\
=\ & \left(\boldsymbol{x}'-\boldsymbol{b}'\boldsymbol{A}^{-1}\right)\boldsymbol{A}\left(\boldsymbol{x}-\boldsymbol{A}^{-1}\boldsymbol{b}\right)-\boldsymbol{b}'\boldsymbol{A}^{-1}\boldsymbol{b}+c \\
=\ & \left(\boldsymbol{x}'\boldsymbol{A}-\boldsymbol{b}'\right)\left(\boldsymbol{x}-\boldsymbol{A}^{-1}\boldsymbol{b}\right)-\boldsymbol{b}'\boldsymbol{A}^{-1}\boldsymbol{b}+c \\
=\ & \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}'\boldsymbol{x}-\boldsymbol{x}'\boldsymbol{b}+\boldsymbol{b}'\boldsymbol{A}^{-1}\boldsymbol{b}-\boldsymbol{b}'\boldsymbol{A}^{-1}\boldsymbol{b}+c \\
=\ & \boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}-2\boldsymbol{b}'\boldsymbol{x}+c
\end{aligned}
$$

Note that the choice of using the term $2\boldsymbol{b}'\boldsymbol{x}$ instead of $\boldsymbol{b}'\boldsymbol{x}$ is purely to simplify the completed square form; the vector of constants $\boldsymbol{b}$ can be arbitrarily scaled.