Self-assess quiz for AMSI ACE Network 2024 course Theory and Methods of Modern Optimisation

Prof. Andrew C. Eberhard
RMIT University

Dr. Scott B. Lindstrom
Curtin Centre for Optimisation and Decision Science

January 10, 2024

Abstract

As necessary background for the course, you will need to have taken undergraduate courses in real analysis, basic vector calculus and linear algebra. It is also desirable to have some familiarity with a computer programming language like Matlab, Julia, Python, etc. In the course, we will be using Python. The following questions will help you assess your readiness for the course.

1 Quiz

Vectors and matrices

Consider the matrices

$$A := \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} \quad \text{and} \quad B := \begin{vmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{vmatrix}.$$

1. Show that A is invertible, or explain why it is not invertible.
2. Show that B is invertible, or explain why it is not invertible.
3. Calculate $A + B$ or explain why it does not exist.
4. Calculate AB or explain why AB does not exist.
5. Calculate BA or explain why BA does not exist.
6. Calculate $B^T A$ or explain why $B^T A$ does not exist.

Real analysis

7. Let $p : \mathbb{R}^n \to \mathbb{R}$ by $p : x \mapsto \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$. Use the Cauchy–Schwarz inequality to show that p is a norm. It is completely fine if you need to look up the definition of the Cauchy–Schwarz inequality.

8. Let $x \in \mathbb{R}^n$ and define $\|x\|_1 := \sum_{i=1}^n |x_i|$ and $\|x\|_2 := \sqrt{\sum_{i=1}^n x_i^2}$. Show that these norms are equivalent.

9. Show that the dot product is an inner product on $\mathbb{R}^n$.

10. Let $(X, \|\cdot\|)$ be a normed space. Suppose that the sequence x_n satisfies $\lim_{n\to\infty} x_n = x$ and also that the sequence y_n satisfies $\lim_{n\to\infty} \|y_n - x_n\| \to 0$. Show that $\lim_{n\to\infty} y_n = x$.

Vector calculus

11. Compute the gradient $\nabla f(x, y)$ where $f(x, y) = x^2 y + y e^{xy}$.

2 Solutions

1. A is invertible, because its column vectors are linearly independent. To verify this, suppose that $x_1 v_1 + x_2 v_2 + x_3 v_3 = 0$ where v_1, v_2, v_3 are the column vectors of A from left to right and $x_1, x_2, x_3 \in \mathbb{R}$. This is just $Ax = 0$, and so we obtain the system

$$\left|\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 10 & 0 \end{array}\right|$$

We can put this in reduced row echelon form as follows:

$$\left|\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right|,$$

which shows that $x = 0$.

2. B cannot be invertible, because it is not a square matrix.

3. $A + B$ does not exist because A and B do not have the same dimensions.

4. We calculate AB by matrix multiplication:

$$AB = \left|\begin{array}{cc} 1\cdot 1 + 2\cdot 3 + 3\cdot 5 & 1\cdot 2 + 2\cdot 4 + 3\cdot 6 \\ 4\cdot 1 + 5\cdot 3 + 6\cdot 5 & 4\cdot 2 + 5\cdot 4 + 6\cdot 6 \\ 7\cdot 1 + 8\cdot 3 + 10\cdot 5 & 7\cdot 2 + 8\cdot 4 + 10\cdot 6 \end{array}\right|.$$

5. BA does not exist because the number of rows of B does not match the number of columns in A.

6. $B^T A$ is computed as follows:

$$B^T A = \left|\begin{array}{ccc} 1\cdot 1 + 3\cdot 2 + 5\cdot 3 & 1\cdot 4 + 3\cdot 5 + 5\cdot 6 & 1\cdot 7 + 3\cdot 8 + 5\cdot 10 \\ 2\cdot 1 + 4\cdot 2 + 6\cdot 3 & 2\cdot 4 + 4\cdot 5 + 6\cdot 6 & 2\cdot 7 + 4\cdot 8 + 6\cdot 10 \end{array}\right|.$$

7. We show p is a norm.

   - **Positive definite:** Notice that $p(x) = 0$ if and only if $x_1 = x_2 = \cdots = x_n = 0$, in which case $x = 0$.
   - **Absolute homogeneity**: Let $\lambda \in \mathbb{R}$. Then

     $$p(\lambda x) = \sqrt{(\lambda x_1)^2 + \cdots + (\lambda x_n)^2} = \sqrt{\lambda^2(x_1^2 + \cdots + x_n^2)} = |\lambda|\sqrt{x_1^2 + \cdots + x_n^2} = |\lambda| p(x).$$

   - **Triangle inquality**: Let $x, y \in \mathbb{R}^n$. From the Cauchy-Schwarz inequality, it holds that

     $$\left(\sum_{i=1}^n x_i y_i\right)^2 \leq \left(\sum_{i=1}^n x_i^2\right)\left(\sum_{i=1}^n y_i^2\right),$$

     or equivalently

     $$\sum_{i=1}^n x_i y_i \leq \sqrt{\sum_{i=1}^n x_i^2} \cdot \sqrt{\sum_{i=1}^n y_i^2} = p(x) \cdot p(y).$$

     Then

     $$\begin{aligned} p(x+y)^2 &= \sum_{i=1}^n (x_i + y_i)^2 \\ &= \sum_{i=1}^n x_i^2 + y_i^2 + 2x_i y_i \\ &= p(x)^2 + p(y)^2 + 2\sum_{i=1}^n x_i y_i \\ \text{(Cauchy-Schwarz)} \quad &\leq p(x)^2 + p(y)^2 + 2p(x)p(y) \\ &= (p(x) + p(y))^2. \end{aligned}$$

8. We must find α and β in $\mathbb{R}$ such that for any $x \in \mathbb{R}^n$:

   $$\alpha\|x\|_1 \leq \|x\|_2 \leq \beta\|x\|_1.$$

   We have that

   $$\|x\|_2^2 = \sum_{i=1}^n x_i^2 \leq \left(\sum_{i=1}^n |x_i|\right)^2 = \|x\|_1^2.$$

   Thus, we can use $\beta = 1$. In fact, this choice is tight. Now, also,

   $$\|x\|_1 = \left(\sum_{i=1}^n |x_i|\right) \leq n \max\{|x_1|, \ldots, |x_n|\} \leq n\|x\|_2.$$

   Thus we can use $\alpha = 1/n$. To get a tight value ($\alpha = 1/\sqrt{n}$), we can use the Cauchy-Schwarz inequality as follows:

   $$\|x\|_1 = \sum_{i=1}^n |x_i| \overset{\text{Cauchy–Schwarz}}{\leq} \sqrt{\sum_{i=1}^n |x_i|^2}\sqrt{\sum_{i=1}^n 1^2} = \sqrt{n}\|x\|_2.$$

9. We show the required properties. Let $x, y, z \in \mathbb{R}^n$. Then

   - $\langle 0, x\rangle = \sum_{i=1}^n 0 \cdot x_i = 0 = \sum_{i=1}^n x_i \cdot 0 = \langle x, 0\rangle$.
   - $\langle x, x\rangle = \sum_{i=1}^n x_i \cdot x_i = \sum_{i=1}^n x_i^2 \geq 0$. Also, the final inequality is only an equality if $x_i = 0$ for all i, whereupon $x = 0$.
   - $\langle x, ay + bz\rangle = \sum_{i=1}^n x_i(ay_i + bz_i) = a\sum_{i=1}^n x_i y_i + b\sum_{i=1}^n x_i z_i = a\langle x, y\rangle + b\langle x, z\rangle$.
   - $\langle x, y\rangle = \sum_{i=1}^n x_i y_i = \sum_{i=1}^n y_i x_i = \langle y, x\rangle$.

   This shows it.

10. By the triangle inequality,

    $$\|y_n - x\| \leq \|y_n - x_n\| + \|x_n - x\|.$$

    Taking the limit as $n \to \infty$, we have that the right hand side approaches zero, which means that y_n converges to x.

11. The gradient consists of the partial derivatives in the variables x and y as follows:

    $$\nabla f(x, y) = \left(2xy + y^2 e^{xy}, x^2 + yxe^{xy} + e^{xy}\right)$$