AMSI summer school 2024
Convex and conic optimisation: pre-enrolment quiz

James Saunderson, Monash University

Linear algebra

Linear algebra is the basic language in which we express ideas in convex optimisation and in which we describe many applications we will see. Fluency in basic manipulations of matrices and vectors in coordinates is expected, as well as being able to reason more abstractly about linear maps and related data (eigenvalues, vector spaces associated with linear maps, coordinate transformations, etc).

1. Let A be an $n \times n$ matrix with real entries. Let the columns of A be $a_1, a_2, \ldots, a_n \in \mathbb{R}^n$. Suppose that $a_1, a_2, \ldots, a_{n-1}$ are linearly independent and $a_n = \sum_{i=1}^{n-1} a_i$.
   (a) What is the dimension of the column space (or image) of A, i.e., $\{Ax \,:\, x \in \mathbb{R}^n\}$?
   (b) What is the dimension of the nullspace (or kernel) of A, i.e., $\{x \in \mathbb{R}^n \,:\, Ax = 0\}$? Find a basis for the nullspace of A.
   (c) What is $\det(A)$?

2. An $n \times n$ matrix with real entries is *symmetric* if $A^\top = A$, i.e., $A_{ij} = A_{ji}$ for all $1 \leq i, j \leq n$. A basic fact about symmetric matrices is that they can be decomposed as $A = U\Lambda U^\top$ where Λ is a diagonal matrix with real diagonal entries and U is an orthogonal matrix, i.e., $U^\top U = UU^\top = I$.
   (a) Let e_i be the ith elementary basis vector (i.e., e_i is zero, except in the ith entry, where it takes value 1). Show that Ue_i is an eigenvector of A. Write down an expression for the corresponding eigenvalue in terms of the entries of Λ.
   (b) Find an expression for $\text{tr}(A^2)$ in terms of the entries of Λ.

3. If A and B are two $n \times m$ matrices, show that $\text{tr}(A^\top B) = \sum_{i=1}^{n} \sum_{j=1}^{m} A_{ij} B_{ij}$.

Multivariable calculus

4. Let $a \in \mathbb{R}^n$ be a non-zero vector and let $f : \mathbb{R}^n \to \mathbb{R}$ be defined by

$$f(x) = \exp(a^\top x) = \exp(\textstyle\sum_{i=1}^{n} a_i x_i).$$

   (a) Find $\nabla f(x)$, the gradient of f at x.
   (b) Find $H_f(x)$, the Hessian of f at x.
   (c) What is the rank of $H_f(x)$?

Analysis/point-set topology

We will be using basic language from real analysis and point-set topology to precisely formulate some of the more subtle aspects of the theory. The following problems check your understanding of open and closed sets, and the notion of infimum.

5. Recall that $S \subseteq \mathbb{R}^n$ is *open* (with respect to the Euclidean topology on $\mathbb{R}^n$) if for every $x_0 \in S$ there exists $\epsilon > 0$ such that $\{x \in \mathbb{R}^n \ : \ \|x - x_0\| < \epsilon\} \subseteq S$. A subset $S \subseteq \mathbb{R}^n$ is *closed* (with respect to the Euclidean topology on $\mathbb{R}^n$) if $\mathbb{R}^n \setminus S$ is open. Decide whether the following sets are open, closed, both open and closed, or neither open nor closed.

   (a) $S = \{(x, y) \in \mathbb{R}^2 \ : \ -1 < x \leq 1, \ -1 \leq y \leq 1\}$

   (b) $S = \{(x, y) \in \mathbb{R}^2 \ : \ -1 \leq x^2 + 2xy - y^2 \leq 0\}$

6. Recall that for $S \subseteq \mathbb{R}^n$:

   - The *interior* of S is the union of all open subsets of $\mathbb{R}^n$ contained in S.
   - The *closure* of S is the intersection of all closed subsets of $\mathbb{R}^n$ containing S.

   For each of the following subsets of $\mathbb{R}^n$, write down its interior and its closure.

   (a) $S = \{(x, y) \in \mathbb{R}^2 \ : \ x + y = 1\}$

   (b) $S = \{(x, y) \in \mathbb{R}^2 \ : \ (x, y) \neq (0, 0)\}$

7. Find $\inf\{e^{-x} \ : \ x \in \mathbb{R}\}$. Is the infimum achieved? (i.e., Does there exist $z \in \mathbb{R}$ such that $e^{-z} = \inf\{e^{-x} \ : \ x \in \mathbb{R}\}$?)

Solutions:

1. (a) The column space is the span of all of the columns of A. This has dimension at least $n-1$ since $a_1, a_2, \ldots, a_{n-1}$ are linearly independent. On the other hand, since $a_n = \sum_{i=1}^{n-1} a_i$, an arbitrary element of the column space can be expressed as

$$Ax = \sum_{i=1}^{n} a_i x_i = \sum_{i=1}^{n-1} a_i(x_i + x_n)$$

and so is an element of the span of $a_1, a_2, \ldots, a_{n-1}$, which has dimension at most $n-1$. Therefore the column space has dimension $n-1$.

(b) Since the dimension of the column space and the dimension of the nullspace sum to n (the rank-nullity theorem), it follows that the nullspace has dimension 1. Since $a_n = \sum_{i=1}^{n-1} a_i$ we know that $x_0 = \begin{bmatrix} 1 & 1 & \cdots & 1 & -1 \end{bmatrix}^\top$ is a non-zero element of the nullspace of A. Since the nullspace is 1-dimensional, x_0 forms a basis for the nullspace of A.

(c) Since A does not have full rank, it is not invertible, and so $\det(A) = 0$.

2. (a) To see that Ue_i is an eigenvector, we check that

$$AUe_i = U\Lambda U^\top U e_i = U\Lambda e_i = U(\Lambda_{ii} e_i) = \Lambda_{ii} U e_i$$

where the first equality is from the eigendecomposition of A, the second is from the fact that U is an orthogonal matrix, and the third holds since Λ is diagonal and e_i is the ith elementary basis vector. We also need to verify that Ue_i is non-zero. To see this we note that $\|Ue_i\|^2 = e_i^\top U^\top U e_i = e_i^\top e_i = 1$ (i.e., e_i is non-zero and orthogonal transformations do not change the Euclidean norm of vectors).

(b) The trace of a square matrix is the sum of its eigenvalues. Since $A^2 = U\Lambda U^\top U \Lambda U^\top = U\Lambda^2 U^\top$ and Λ^2 is diagonal, with diagonal entries Λ_{ii}^2, it follows (essentially from the previous problem) that the eigenvalues of A^2 are Λ_i^2 for $i = 1, 2, \ldots, n$. Therefore, the sum of the eigenvalues of A^2 is $\sum_{i=1}^{n} \Lambda_{ii}^2$. (There are many other ways to do this problem. Another approach is to use the fact that tr$(ABC) =$ tr(BCA) for any matrices of compatible sizs, and so tr$(A^2) =$ tr$(U\Lambda U^\top U \Lambda U^\top) =$ tr$(U\Lambda^2 U^\top) =$ tr$(\Lambda^2 U^\top U) =$ tr$(\Lambda^2) = \sum_{i=1}^{n} \Lambda_{ii}^2$.)

3. The (j, j) entry of $A^\top B$ is $[A^\top B]_{jj} = \sum_{i=1}^{n} A_{ij} B_{ij}$. (Another way to think of this is that it is the dot product of the jth column of B and the jth row of $A^\top$, which is the jth column of A.) The trace is the sum of the diagonal entries, so is

$$\text{tr}(A^\top B) = \sum_{j=1}^{m} [A^\top B]_{jj} = \sum_{j=1}^{m} \sum_{i=1}^{n} A_{ij} B_{ij}.$$

4. Let $f(x) = \exp(a^\top x)$.

(a) Computing the partial derivative with respect to the ith variable gives $\frac{\partial f}{\partial x_i} = a_i \exp(a^\top x)$. Therefore $\nabla f(x) = \exp(a^\top x) a$.

(b) Computing the second partial derivative gives $\frac{\partial^2 f}{\partial x_i \partial x_j} = a_i a_j \exp(a^\top x)$. Therefore $H_f(x) = \exp(a^\top x) a a^\top$.

(c) Since $\exp(a^\top x) > 0$ for all x, and $a \neq 0$, the rank of $H_f(x)$ is one, because $aa^\top$ has rank one.

(What happens, in general, for functions $f : \mathbb{R}^n \to \mathbb{R}$ of the form $f(x) = g(a^\top x)$ where g is a smooth univariate function?)

5. (a) S is neither open nor closed. To see that it is not open, we note that $(x, y) = (1, 1)$ is not an interior point of S. To see that it is not closed, we see that $(-1, 0)$ is not an interior point of the complement of S.

   (b) S is closed. This is because $f(x, y) = x^2 + 2xy - y^2$ is continuous, so $S = \{(x, y) : -1 \leq f(x, y) \leq 0\} = f^{-1}([-1, 0])$ is the preimage of a closed set under a continuous mapping, and so is a closed set.

6. (a) The interior of S is the empty set, since no point of S has an open neighborhood contained in S. Since S is closed, S equals its closure.

   (b) Since S is open, the interior of S is S itself. One way to understand the closure of S is to consider the complement $\overline{S} = \mathbb{R}^2 \setminus S = \{(0, 0)\}$. The interior of $\overline{S} = \emptyset$. Now the closure of S is the complement of the interior of $\overline{S}$. Therefore the closure of S is $\mathbb{R}^2$.

7. The infimum is 0. First note that 0 is a lower bound on $S = \{e^{-x} \ : \ x \in \mathbb{R}\}$ because $0 \leq e^{-x}$ for all $x \in \mathbb{R}$. To see that this is the greatest lower bound, we argue by contradiction. Suppose there was some $\epsilon > 0$ that was a lower bound on S. If $x > \log_e(1/\epsilon)$ then $e^{-x} < \epsilon$, contradicting our assumption that S is a lower bound. So no positive real number can be a lower bound on $\{e^{-x} \ : \ x \in \mathbb{R}\}$.

   The infimum is not achieved, since there is no real number z such that $e^{-z} = 0$.