

# 18.06 - Review for final exam

Sam Turton

May 14, 2019

*This is a non-exhaustive list of material we have covered this semester. It is posted on github and I will update it as I think of more things. Please email me (seturton@mit.edu) if you think there are sections I should add!*

## 1 Vectors

- The *length* (or *magnitude*) of a vector  $\mathbf{v} \in \mathbb{R}^n$  is written  $\|\mathbf{v}\|$ . It is given by the following formula:

$$\|\mathbf{v}\|^2 = v_1^2 + v_2^2 + \dots + v_n^2 = \sum_{i=1}^n v_i^2$$

- A *unit vector*  $\mathbf{n}$  is a vector with length  $\|\mathbf{n}\| = 1$ .
- Suppose you have two  $n$ -dimensional vectors,  $\mathbf{u}$  and  $\mathbf{v}$ :

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

Then the *dot product* (or *inner product*) of these two vectors,  $\mathbf{u} \cdot \mathbf{v}$ , is given by the following formula:

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n$$

- The *angle* between two vectors  $\mathbf{u}$  and  $\mathbf{v}$  is given by the following formula:

$$\cos \theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}$$

- We say that a nonzero vector  $\mathbf{u}$  is *parallel* to a nonzero vector  $\mathbf{v}$  if  $\boxed{\mathbf{u} = a\mathbf{v}}$  for some scalar  $a \neq 0$ . We sometimes say that  $\mathbf{u}$  and  $\mathbf{v}$  are in the same direction.
- We say that a vector  $\mathbf{u}$  is *perpendicular*, or *orthogonal*, to a vector  $\mathbf{v}$  if  $\boxed{\mathbf{u} \cdot \mathbf{v} = 0}$ .

## 2 Matrices

- A *matrix* is an  $m \times n$  array of numbers. An  $m \times n$  matrix has  $m$  *rows* and  $n$  *columns*. A matrix is *square* if  $m = n$ .
- Suppose  $A$  is a  $m \times n$  matrix and  $B$  is a  $p \times q$  matrix. We can only multiply these matrices if the dimensions make sense. We can multiply  $AB$  only if  $n = p$ ; we can multiply  $BA$  only if  $m = q$ .
- Suppose  $A$  and  $B$  are two  $n \times n$  square matrices. In general  $\boxed{AB \neq BA}$ . Matrix multiplication does not *commute*.
- A *diagonal* matrix is a matrix which only has entries along its diagonal.
- We can compute the product of three matrices  $ABC$  either as  $(AB)C$  (multiply  $AB$  and then multiply on the right by  $C$ ), or as  $A(BC)$  (multiply  $BC$  and then multiply on the left by  $A$ ). Matrix multiplication is *associative*.
- Matrix multiplication is *distributive*. This means that  $(A + B)C = AC + BC$ .
- The *transpose* of a matrix  $A$  is denoted by  $A^T$ . The transpose is the matrix formed by taking the columns of  $A$  and making them the rows of  $A^T$ .
- If  $A$  has components  $a_{ij}$ , then  $A^T$  has components  $a_{ji}$ .
- If  $A$  is  $m \times n$ , then  $A^T$  is  $n \times m$ .
- $A^T A$  and  $AA^T$  can always be computed. They are both square matrices, even if  $A$  is rectangular. If  $A$  is  $m \times n$ , then  $A^T A$  is  $m \times m$  and  $AA^T$  is  $n \times n$ .
- $\boxed{(AB)^T = B^T A^T}$ .
- $\boxed{(AB)^{-1} = B^{-1} A^{-1}}$  provided  $A^{-1}$  and  $B^{-1}$  exist.
- If  $A$  is a square matrix and  $A^{-1}$  exists, then  $\boxed{(A^T)^{-1} = (A^{-1})^T}$ .
- A  $n \times n$  square matrix  $Q$  is *orthogonal* if  $\boxed{Q^T Q = I}$ . A rectangular matrix for which  $Q^T Q = I$  we will usually refer to as a *tall skinny orthogonal* matrix.
- An equivalent definition: suppose a matrix  $Q$  has  $n$  columns given by the vectors  $\mathbf{q}_1, \dots, \mathbf{q}_n$ . Then  $Q$  is orthogonal if the column vectors are orthonormal. This means that  $q_i^T q_j = 0$  for  $i \neq j$ , and  $\|q_i\| = 1$ .
- If  $Q$  is square and orthogonal, then we also have that  $QQ^T = I$ , and so  $Q^{-1} = Q^T$ .
- If  $Q$  obey  $Q^T Q = I$ , but is not square, then  $QQ^T \neq I$  generally!

### 3 Vector spaces, independence and bases

- A **vector subspace**,  $X$ , is a set of vectors that satisfies the following requirement: *If  $v$  and  $w$  are vectors in  $X$ , and  $c$  is any scalar, then*

1.  $v + w$  is in  $X$

2.  $cv$  is in  $X$

- A set of vectors  $v_1, v_2, \dots, v_n$  is linearly independent if

$$a_1v_1 + a_2v_2 + \dots + a_nv_n = 0 \iff a_1 = a_2 = \dots = a_n = 0$$

- A set of vector  $v_1, v_2, \dots, v_n$  *spans* a vector space if every element of the space can be written as a linear combination of these  $n$  vectors.
- A set of vector  $v_1, v_2, \dots, v_n$  is a *basis* if the vectors are linearly independent and span the space.
- The *dimension* of a vector space is the number of elements in a basis for the vector space (this number is well-defined).
- The *standard basis* for  $\mathbb{R}^n$  is the set of vectors  $e_1, e_2, \dots, e_n$ . The vectors  $e_i$  have a 1 in their  $i$ -th component, and every other component is zero.

## 4 Linear systems

- We say that an  $m \times n$  matrix  $A$  has full column rank if  $r = n$ , and full row rank if  $r = m$ . The rank is always less than or equal to the minimum of  $m$  and  $n$ .
- There are four possibilities for a linear system  $Ax = b$ :
  1.  $Ax = b$  has a unique solution for all  $b$ . This can only happen if  $r = m = n$  (full column and row rank)
  2.  $Ax = b$  has either a unique solution or no solution, depending on  $b$ . This can only happen if  $r = n$ , but  $m > n$  (full column rank, not full row rank).
  3.  $Ax = b$  has infinitely many solutions for all  $b$ . This can only happen if  $r = m$ , but  $n > m$  (full row rank, not full column rank).
  4.  $Ax = b$  has either no solution or infinitely many solutions, depending on  $b$ . This can only happen if  $r < \min(m, n)$  (neither full column nor row rank).

## 5 The four fundamental subspaces

- Let  $A = U\Sigma V^T$  be the full SVD for the  $m \times n$  matrix  $A$ .
- The **column space**  $C(A)$  is the set of vectors in  $\mathbb{R}^m$  which are spanned by the columns of  $A$ . Equivalently, it is all vectors  $u \in \mathbb{R}^m$  that can be written  $u = Ax$  for any  $x \in \mathbb{R}^n$ . The first  $r$  columns of  $U$  are a basis for  $C(A)$ .
- The **row space**  $\text{row}(A) = C(A^T)$  is the set of vectors in  $\mathbb{R}^n$  which are spanned by the rows of  $A$ . Equivalently, it is all vectors  $v \in \mathbb{R}^n$  that can be written  $v = A^T x$  for any  $x \in \mathbb{R}^m$ . The first  $r$  columns of  $V$  are a basis for  $C(A)$ .
- The **null space**  $N(A)$  is the set of vectors in  $\mathbb{R}^n$  which satisfy  $Ax = 0$ . The last  $n - r$  columns of  $V$  are a basis for  $N(A)$ .
- The **left null space**  $N(A^T)$  is the set of vectors in  $\mathbb{R}^m$  which satisfy  $A^T x = 0$ . The last  $m - r$  columns of  $U$  are a basis for  $N(A^T)$ .
- Vectors in  $C(A)$  are orthogonal to vectors in  $N(A^T)$ . Vectors in  $C(A^T)$  are orthogonal to vectors in  $N(A)$ .

## 6 Linear transformations

- A linear transformation  $T$  takes vectors  $v \in V$  to vectors  $T(v) \in W$ , where  $V$  and  $W$  are vector spaces. Linearity requires

$$T(cv_1 + dv_2) = cT(v_1) + dT(v_2)$$

for any vectors  $v_1, v_2 \in V$  and scalars  $c, d \in \mathbb{R}$ . Note that  $T(0) = 0$  necessarily.

- A linear transformation is uniquely defined by its action on a basis, i.e. if  $\{v_1, \dots, v_n\}$  is a basis for a vector space, then any vector can be written as  $v = c_1v_1 + \dots + c_nv_n$ . Therefore

$$T(v) = c_1T(v_1) + \dots + c_nT(v_n),$$

and so knowing  $T(v_1), \dots, T(v_n)$  allows us to determine  $T(v)$  for any  $v$  in the vector space.

- A linear transformation  $T(v)$  can be described by a matrix, i.e.  $T(v) = Av$ . Column  $j$  in the matrix  $A$  comes from applying  $T$  to the basis vector  $v_j$ , where  $\{v_1, v_2, \dots, v_n\}$  is a basis for the vector space  $V$ .

## 7 Determinants, cofactors and inverses

Let  $A$  be a square  $n \times n$  matrix.

- The *determinant* of  $A$  is the unique function  $\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$  for which
  - $\det I = 1$  for the identity matrix of any dimension
  - $\det A$  changes sign when any two rows of  $A$  are interchanged.
  - $\det A$  is a linear transformation of each row of  $A$ .
- The determinant of a  $2 \times 2$  matrix has a simple formula:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \implies \det A = ad - bc$$

- Two very useful rules for determinants:

$$\det A^T = \det A, \quad \det AB = \det A \det B$$

- The *cofactor*  $C_{ij}$  is defined by

$$C_{ij} = (-1)^{i+j} \det M_{ij}$$

where  $M_{ij}$  is the  $(n-1) \times (n-1)$  matrix obtained by removing the  $i$ th row and  $j$ th column from  $A$ .

- **Compute determinant by cofactors:** For any  $1 \leq i \leq n$ ,

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \dots + a_{in}C_{in}.$$

- If  $A$  is invertible, then

$$A^{-1} = \frac{1}{\det A} C^T.$$

In terms of entries,  $(A^{-1})_{ij} = \frac{C_{ji}}{\det A}$ .

- **Cramer's Rule:** If  $A$  is invertible and  $Ax = b$ , then

$$x_i = \frac{\det B_i}{\det A},$$

where  $B_i$  is the matrix  $A$  with the  $i$ th column replaced by  $b$ .

## 8 Eigenvalues, eigenvectors and diagonalisation

- If we have that

$$Ax = \lambda x, \quad x \neq 0,$$

then  $\lambda$  is an **eigenvalue** of  $A$ , and  $x$  is the corresponding **eigenvector**.

- The eigenvalues are solutions of the  $n$ -th degree polynomial equation

$$\det(A - \lambda I) = 0.$$

This is known as the *characteristic equation*.

- Warning: The eigenvalues and eigenvectors need not be real, even if all the entries of  $A$  are real !
- $A$  is invertible if, and only if, 0 is **not** an eigenvalue.
- Powers of  $A$  have the same eigenvectors as  $A$ ; the corresponding eigenvalues are raised to the same power, i.e.

$$Ax = \lambda x \implies A^n x = \lambda^n x$$

- The eigenvalues of a triangular matrix are on the diagonal.
- The determinant of  $A$  is equal to the product of all of its eigenvalues.
- The trace of  $A$  is equal to the sum of its eigenvalues.
- **Diagonalisation:** Suppose  $A$  has  $n$ -independent eigenvectors  $x_1, x_2, \dots, x_n$ , with corresponding eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_n$ . Let  $X$  be a matrix whose columns are the eigenvectors in this order, and  $\Lambda$  be a diagonal matrix with the corresponding eigenvalues along its diagonal. Then we can write

$$A = X\Lambda X^{-1}$$

- Diagonalisation is always possible if  $A$  has  $n$  distinct eigenvalues; it may or may not be possible if it has repeated eigenvalues.

## 9 Positive definiteness and Markov matrices

- A symmetric matrix  $A$  is positive definite if, and only if, any of the following are true:
  1. All of the eigenvalues of  $A$  are strictly positive
  2. All upper left determinants are strictly positive
  3.  $x^T A x > 0$  for all nonzero vectors  $x$
  4.  $A = B^T B$ , where  $B$  has independent columns, but  $B$  is not necessarily square.
- A *Markov matrix* has non-negative entries and the elements of each column sum to 1. It always has one eigenvalue equal to 1, and every other eigenvalue has absolute value less than or equal to 1.
- A *positive Markov matrix* has strictly positive entries and the elements of each column sum to 1. It always has one eigenvalue equal to 1, and every other eigenvalue has absolute value strictly less than 1.
- The eigenvector of a Markov matrix corresponding to  $\lambda = 1$  is called the *steady state* vector.