Recitation notes

May 30, 2008

1 Linear algebra definitions, and facts

Most of the statements here will be over $\mathbb{C}$, but of course they may be restricted to $\mathbb{R}$.

1.1 Definitions

1.1.1 Matrix operations

- The *inverse* of a square matrix A (if it exists) is a matrix A^{-1} which satisfies $A^{-1}A = I$
- The *complex conjugate* of A is $\overline{A}$, the matrix which results from taking the complex conjugate of every element of A
- The *Hermitian adjoint* of A is $A^* = \overline{A}^T = \overline{A^T}$
- The *trace* of A is $\mathrm{tr}A$, the sum of the diagonal elements of A

1.1.2 Matrix types

A matrix is:

- *nonsingular* iff A has an inverse
- *normal* iff $A^*A = AA^*$
- *Hermitian* iff $A = A^*$
- *unitary* iff $A^*A = I$
- *symmetric* iff $A = A^T$
- *orthogonal* iff $A^TA = I$
- *positive definite* iff A is Hermitian, and all eigenvalues of A are positive reals
- *positive semi-definite* iff A is Hermitian, and all eigenvalues of A are nonnegative reals

Whenever one speaks of symmetric or orthogonal matrices, one is generally speaking of real matrices with this property, in which case symmetric is equivalent to Hermitian, and orthogonal is equivalent to unitary.

1.1.3 Eigenvalues and eigenvectors

- If $Ax = \lambda x$ for some $\lambda \in \mathbb{C}$, then x is said to be an *eigenvector* of A, with associated *eigenvalue* λ. Note that if $Ax = \lambda x$, then $A(cx) = \lambda(cx)$ for any $c \in \mathbb{C}$–that is, the scale of the eigenvector is irrelevant. Hence, one generally gives eigenvectors with unit l^2 norm ($x^* x = 1$).
- We may have that there are multiple eigenvectors with the same eigenvalue. For example, the identity matrix has every vector as an eigenvector, all with associated eigenvalue 1. In this case, we say that the matrix has eigenvalue λ with *multiplicity* k, where k is the dimension of the space spanned by the eigenvectors.
- If $A \in \mathbb{C}^{n\times n}$, and there exist n orthonormal eigenvectors $v_i : i \in \{1, 2, \ldots, n\}$ of A (so that $v_i^* v_j = \delta_{i,j}$), then we say that these eigenvectors form an *eigenbasis* of A. By letting $U \in \mathbb{C}^{n\times n}$ contain each of these eigenvectors in one of its columns (so $U_{i,:} = v_i$), we will have that $U^* U = I$ by orthonormality, so that U is unitary. Furthermore, if we let Λ be the diagonal matrix, where the ith diagonal entry contains the associated eigenvalue of the ith column of U, then we'll have that:

  $$(U\Lambda U^*)\, U_{i,:} = \lambda_i U_{i,:}$$

  Showing that $U_{i,:}$ is an eigenvector of $U\Lambda U^*$ with associated eigenvalue λ_i. Since $U_{i,:}$ spans $\mathbb{C}^n$, this shows that $U\Lambda U^*$ acts on every vector on $\mathbb{C}^n$ in exactly the same way as A, so that $A = U\Lambda U^*$. One says that such an A is *unitarily diagonalizable*.

1.2 Facts

- If A is nonsingular, then A^{-1} is unique, and is both a left and right inverse (so $AA^{-1} = A^{-1}A = I$)
- If A is Hermitian, then all eigenvalues of A are real
- If A is Hermitian or unitary, then it is normal. In particular, if A is real symmetric or orthogonal, then it is normal.
- Every normal matrix has an eigenbasis.

The fact that a Hermitian matrix has real eigenvalues, and an orthonormal eigenbasis, is known as the *spectral theorem*.

2 Simple facts with proof

2.1 The eigenvalues of $tI + A$ are the eigenvalues of A plus t

Suppose that v is an eigenvector of A with associated eigenvalue λ. Then:

$$\begin{aligned}(tI + A)\, v &= tv + \lambda v \\ &= (t + \lambda)\, v\end{aligned}$$

So v is an eigenvector of $tI + A$, with associated eigenvalue $t + \lambda$.

2.2 If v is an eV of A with associated ev λ, then $v^T Av = \lambda \|v\|_2^2$

By definition, $Av = \lambda v$, so:

$$\begin{aligned}v^T Av &= \lambda v^T v \\ &= \lambda \|v\|_2^2\end{aligned}$$

2.3 If A is normal with ev,eVs λ_i, v_i, then we may write any $x = \sum_{i=1}^n \alpha_i v_i$, and $Ax = \sum_{i=1}^n \lambda_i \alpha_i v_i$

Suppose that $A \in \mathbb{C}^{n \times n}$ is normal. Then it has an orthonormal eigenbasis, so that we may write any $x \in \mathbb{C}^n$ as $x = \sum_{i=1}^n \alpha_i v_i$, where v_i is the (normalized) ith eigenvector of A, with associated eigenvalue λ_i. Hence:

$$\begin{aligned} Ax &= \sum_{i=1}^{n} \alpha_i A v_i \\ &= \sum_{i=1}^{n} \lambda_i \alpha_i v_i \end{aligned}$$

2.4 A Hermitian matrix A is PSD iff $x^* A x \geq 0$ for all x

2.4.1 $\Rightarrow$

If A is PSD, then we may write any x in the form $x = \sum_{i=1}^n \alpha_i v_i$, so that:

$$\begin{aligned} x^* A x &= \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{\alpha_i} \alpha_j v_i^* A v_j \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_j \overline{\alpha_i} \alpha_j v_i^* v_j \end{aligned}$$

Since $\{v_i\}$ is an orthonormal basis, $v_i^* v_j = \delta_{i,j}$, so:

$$x^* A x = \sum_{i=1}^{n} \lambda_i \left|\alpha_i\right|^2$$

And since A is positive semi-definite, $\lambda_i \geq 0$ for all i, so $x^* A x \geq 0$.

2.4.2 $\Leftarrow$

Since A is Hermitian, it has an orthonormal eigenbasis $\{v_i\}$. By assumption, $v_i^* A v_i \geq 0$, so:

$$\begin{aligned} 0 &\leq v_i^* A v_i \\ &\leq v_i^* \lambda_i v_i \\ &\leq \lambda_i \left|v_i\right|^2 \\ &\leq \lambda_i \end{aligned}$$

Showing that A is positive semi-definite.

2.5 A Hermitian matrix A is PD iff $x^* A x > 0$ for all x

The proof of this fact is identical to the proof of the previous fact, with $\geq$ replaced by $>$, and $\leq$ replaced by $<$.

2.6 If A is PSD, then its diagonal elements are nonnegative reals

Since A is PSD, $x^* A x \geq 0$ for all x. In particular, $a_{ii} = e_i^* A e_i \geq 0$, where e_i is the ith standard unit basis vector (the vector with the ith element 1, all others 0), showing that the diagonal elements of A are nonnegative reals.

2.7 If A is PD, then its diagonal elements are positive reals

The proof of this fact is identical to the proof of the previous fact, with $\geq$ replaced by $>$.

2.8 $A \cdot B = \mathbf{tr}\left(A^T B\right)$

Suppose that $A, B \in \mathbb{C}^{n \times n}$. By definition:

$$\begin{aligned} A \cdot B &= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ij} \\ &= \sum_{i=1}^{n} A_{i,:}^T B_{i,:} \\ &= \operatorname{tr}\left(A^T B\right) \end{aligned}$$

2.9 $\mathbf{tr}\,(ABC) = \mathbf{tr}\,(CAB)$

Suppose that $A \in \mathbb{C}^{n_1 \times n_2}$, $B \in \mathbb{C}^{n_2 \times n_3}$, and $C \in \mathbb{C}^{n_3 \times n_1}$. By definition:

$$\begin{aligned} \operatorname{tr}(ABC) &= \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} a_{ij} b_{jk} c_{ki} \\ &= \sum_{k=1}^{n_3} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} c_{ki} a_{ij} b_{jk} \\ &= \operatorname{tr}(CAB) \end{aligned}$$

This result immediately gives us the *rotation rule* for the trace:

$$\begin{aligned} \operatorname{tr}(A_1 \cdots A_k) &= \operatorname{tr}(A_k A_1 \cdots A_{k-1}) \\ &= \operatorname{tr}(A_{k-1} A_k A_1 \cdots A_{k-2}) \\ &\cdots \\ &= \operatorname{tr}(A_3 \cdots A_k A_1 A_2) \\ &= \operatorname{tr}(A_2 \cdots A_k A_1) \end{aligned}$$

2.10 If B is PSD, then $C^* B C$ is PSD

First, note that since B is Hermitian, $C^* B C$ is Hermitian. Furthermore, for any x:

$$x^* C^* B C x = (Cx)^* B (Cx)$$

Since B is positive semi-definite, $y^* B y \geq 0$ for all y, so letting $y = Cx$ gives that:

$$x^* C^* B C x \geq 0$$

the above holding for all x. Hence, $C^* B C$ is positive semi-definite.

In particular, note that this result gives that $C^* I C = C^* C$ is positive semi-definite for every C.

2.11 A Hermitian matrix A is PSD iff $A \cdot B \geq 0$ for all PSD B

2.11.1 $\Rightarrow$

Suppose that A, B are positive semi-definite. Since A is positive semi-definite, it is unitarily diagonalizable, so $A = U \Lambda U^*$, and the eigenvalues (the diagonal elements of Λ) are nonnegative. Letting $\Lambda^{\frac{1}{2}}$ be the matrix which results from taking the square roots of the diagonal elements of Λ, we have that if $C = U \Lambda^{\frac{1}{2}} U^*$, then $CC^* = U \Lambda^{\frac{1}{2}} U^* U \Lambda^{\frac{1}{2}} U^* = U \Lambda^{\frac{1}{2}} \Lambda^{\frac{1}{2}} U^* = A$, so that:

$$\begin{aligned} A \cdot B &= \operatorname{tr}(A^* B) \\ &= \operatorname{tr}(CC^* B) \\ &= \operatorname{tr}(C^* BC) \end{aligned}$$

Since $C^* BC$ is positive semi-definite, $\operatorname{tr}(C^* BC) \geq 0$ (the diagonal elements of a positive semi-definite matrix are nonnegative). Hence:

$$A \cdot B \geq 0$$

2.11.2 $\Leftarrow$

Suppose that A is Hermitian, and that $A \cdot B \geq 0$ for every positive semi-definite B. Let C be the matrix which contains the vector x in its first column, with all other columns being zero. Then $B = CC^*$ is positive semi-definite, and:

$$\begin{aligned} 0 &\leq A \cdot B \\ &\leq \operatorname{tr}(A^* B) \\ &\leq \operatorname{tr}(A^* CC^*) \\ &\leq \operatorname{tr}(C^* A^* C) \\ &\leq x^* A x \end{aligned}$$

The above holding for every x. Hence, A is positive semi-definite.

2.12 If A is PD, B is PSD, and $A \cdot B = 0$, then $B = 0$

Since A is positive definite, it is unitarily diagonalizable, so $A = U \Lambda U^*$, and the eigenvalues (the diagonal elements of Λ) are positive. Letting $\Lambda^{\frac{1}{2}}$ be the matrix which results from taking the square roots of the diagonal elements of Λ, we have that if $C = U \Lambda^{\frac{1}{2}} U^*$, then $CC^* = U \Lambda^{\frac{1}{2}} U^* U \Lambda^{\frac{1}{2}} U^* = U \Lambda^{\frac{1}{2}} \Lambda^{\frac{1}{2}} U^* = A$, so that:

$$\begin{aligned} 0 &= A \cdot B \\ &= \operatorname{tr}(A^* B) \\ &= \operatorname{tr}(CC^* B) \\ &= \operatorname{tr}(C^* BC) \\ &= \operatorname{tr}\left(U \Lambda^{\frac{1}{2}} U^* B U \Lambda^{\frac{1}{2}} U^*\right) \\ &= \operatorname{tr}\left(\Lambda^{\frac{1}{2}} U^* B U \Lambda^{\frac{1}{2}}\right) \\ &= \operatorname{tr}(\Lambda U^* B U) \end{aligned}$$

Since U^*BU is positive semi-definite, we must have that the diagonal elements of U^*BU are nonnegative. Likewise, the diagonal elements of Λ are positive. Hence, the diagonal elements of ΛU^*BU are all nonnegative, so that since they sum to 0, they must all be 0. Since the diagonal elements of Λ are positive, this implies that all of the diagonal elements of U^*BU are 0, so in particular their sum is 0:

$$\begin{aligned} 0 &= \operatorname{tr}(U^*BU) \\ &= \operatorname{tr}B \end{aligned}$$

Since B is positive semi-definite, it is unitarily diagonalizable, so $B = V\Gamma V^*$, with V unitary and Γ the diagonal matrix containing the eigenvalues of B:

$$\begin{aligned} 0 &= \operatorname{tr}(V^*\Gamma V) \\ &= \operatorname{tr}\Gamma \end{aligned}$$

Since the eigenvalues of B are nonnegative, this shows that the eigenvalues are in fact all identically 0, so $B = 0$.