Linear algebra notes: matrices

Stephen Becker

Sept 2021 for APPM 5650, last updated September 4, 2021

Many books have condensed reviews of linear algebra in their appendix; and of course there are plenty of good books on linear algebra. These notes are intended to be an abbreviated introduction to some basic *matrix* facts that sometimes are missed by classes that are either theoretical (basis-free, infinite dimensions, etc.) or too practical (focus on algorithms). We also pay particular attention to the fields. Source code at www.overleaf.com/read/yprqvktvsxgb.

Good references:

1. *Matrix Analysis*, Horn and Johnson; 1st edition 1990 or 2nd edition 2012. This, and its sequel *Topics in Matrix Analysis*, is full of useful facts that are hard to find elsewhere. My citations to "H&J" refer to Horn and Johnson 1st edition.
2. *Matrix Theory*, Joel Franklin; 1968, reprinted 2000 by Dover. Short and sweet.

We write $\mathbb{R}^{m\times n}$ to denote real $m \times n$ matrices, and $\mathbb{C}^{m\times n}$ to denote complex $m \times n$ matrices. We typically work with $\mathbb{C}^{m\times n}$ (or $\mathbb{C}^{n\times n}$ if square) since this includes $\mathbb{R}^{m\times n}$ as a special case.

Chapter 1 in H&J

We assume the reader has some linear algebra background, such as knowing what eigenvectors are, but is rusty and/or may not know all the details of these results. Starting with eigenvalues of a matrix $A \in \mathbb{C}^{n\times n}$, these are the roots of the characteristic equation $p(\lambda) = \det(A - \lambda I)$. Since p is a degree n polynomial, hence it is guaranteed to have n complex roots (eigenvalues). These are not guaranteed to be real, even if $A \in \mathbb{R}^{n\times n}$. Note the **Cayley-Hamilton** theorem which says that $p(A) = 0$.

A square matrix $A \in \mathbb{C}^{n\times n}$ is **diagonalizable** if it is similar to a diagonal matrix, meaning $A = VDV^{-1}$ where D is diagonal and V is nonsingular. If so, then the columns of V are eigenvectors. Not all matrices are diagonalizable; e.g., $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is not diagonalizable. A sufficient condition to be diagonalizable is if all eigenvalues are distinct. More generally, a matrix is diagonalizable if the algebraic multiplicity of its eigenvalues (the repeated roots in p) match the geometric multiplicity (the dimension spanned by the corresponding eigenvectors).

If A is diagonalizable and real-valued, it need not have real eigenvalues. If A is real and an eigenvalue is complex, the corresponding eigenvector must be complex too. If A is real and only has real eigenvalues, the eigenvectors could possibly still be complex. For example, take the identity matrix, which is real-valued. Any nonsingular matrix S is a set of eigenvectors, so this can include complex-valued eigenvectors.

Chapter 2 in H&J

A matrix $U \in \mathbb{C}^{n\times n}$ is **unitary** if $U^*U = I$. A real matrix $Q \in \mathbb{R}^{n\times n}$ such that $Q^TQ = I$ is **real orthogonal** — note that it is a special case of unitary. A *complex* matrix $Q \in \mathbb{C}^{n\times n}$ such that Q^TQ is called **orthogonal**, but in this case it is not a special case of unitary since then $Q^T \neq Q^*$, and these matrices generally do not have nice properties. Many people mean "real orthogonal" when they speak of "orthogonal", which is the convention we'll follow in these notes (though we may try to use "real orthogonal" in order to be very clear).

The following are equivalent: (a) U is unitary; (b) U is nonsingular and $U^* = U^{-1}$; (c) $UU^* = I$; (d) U^* is unitary; (e) the columns of U form an orthonormal set; (f) the rows of U form an orthonormal set; (g) for all $x \in \mathbb{C}^n$, $\|Ux\|_2 = \|x\|_2$. The eigenvalues of unitary matrices need not be real (e.g., a rotation matrix like $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$), but all eigenvalues have magnitude 1, and hence the determinant is also magnitude 1 since the determinant is the product of the eigenvalues; for real orthogonal matrices, the determinant is also real (any complex eigenvalues come in complex conjugate pairs and so multiply to a real number), hence it is ± 1.

Schur's unitary triangularization theorem

Theorem (Schur, 2.3.1 in H&J)**.** *Given $A \in \mathbb{C}^{n\times n}$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, there is a* unitary *matrix $U \in \mathbb{C}^{n\times n}$ such that*

$$A = UTU^*$$

where T is upper triangular with eigenvalues on the diagonal. If also $A \in \mathbb{R}^{n\times n}$ and *all the eigenvalues are real, then U may be chosen to be* real orthogonal *(this also implies T is real).*

Both T and U are not necessarily unique. A variant that applies to *all* real matrices, not just those with real eigenvalues, is the following:

Theorem (2.3.4 in H&J)**.** *Given $A \in \mathbb{R}^{n\times n}$, there is a* real orthogonal *matrix $Q \in \mathbb{R}^{n\times n}$ such that*

$$A = QTQ^T$$

where T is almost upper triangular, in the sense that its diagonal is either scalars (for real eigenvalues) or a real but 2×2 block (corresponding to complex eigenvalues).

One implication of Schur's theorem is that the set of matrices with n distinct eigenvalues are dense in $\mathbb{C}^{n\times n}$. In particular, diagonalizable matrices are dense in $\mathbb{C}^{n\times n}$.

Normal matrices A square matrix $A \in \mathbb{C}^{n\times n}$ is *normal* if $A^*A = AA^*$; this includes unitary and Hermitian (aka self-adjoint, $A = A^*$) matrices as special cases. Note that a real Hermitian matrix is symmetric; as in the case of "orthogonal", a complex matrix is symmetric if $A = A^T$ but this is not that useful of a property since it is not Hermitian. Usually the term "symmetric" means "real symmetric".

Theorem (2.5.4, 2.5.6 and 2.5.8 in H&J)**.** *A matrix $A \in \mathbb{C}^{n\times n}$ is unitarily diagonalizable, meaning $A = QDQ^*$ with Q unitary and D diagonal, if and only if A is* normal. *If $A \in \mathbb{R}^{n\times n}$, then $A = QDQ^T$ with Q orthogonal and D diagonal if and only if A is normal. Whether A is real or complex, if it is Hermitian, then the eigenvalues are real, i.e., $D_{ii} \in \mathbb{R}$.*

QR factorization We now discuss rectangular matrices.

Theorem (QR factorization, 2.6.1 in H&J)**.** *If $A \in \mathbb{C}^{m\times n}$ and $m \ge n$, there is a matrix $Q \in \mathbb{C}^{m\times n}$ with orthonormal columns, and an upper triangular matrix $R \in \mathbb{C}^{n\times n}$ such that*

$$A = QR.$$

If $m = n$ then Q is unitary. If $m = n$ and A is nonsingular, then R may be chosen so that all its diagonal entries are positive, and in this event, both Q and R are unique. If $A \in \mathbb{R}^{m\times n}$ then both Q and R may be taken to be real valued.

In general, if $m \ge n$, then R is nonsingular if A is full-rank.

From the QR, you can show that any positive semi-definite matrix $B \in \mathbb{C}^{n\times n}$ (meaning Hermitian and all eigenvalues are non-negative) has a **Cholesky** factorization $B = LL^*$ with $L \in \mathbb{C}^{n\times n}$ lower-triangular. The factorization is unique if B is positive definite (i.e., positive semi-definite and invertible). A similar factorization for Hermitian matrices (not necessarily positive semi-definite) is the "LDL" factorization, where $B = LDL^*$ with L lower triangular and D diagonal. A similar factorization for general matrices is the "LU" factorization created as a byproduct of Gaussian Elimination, where $B = LU$ with L lower triangular and U upper triangular.

Chapter 3 in H&J

Theorem (Jordan Canonical Form, 3.1.11 in H&J)**.** *Let $A \in \mathbb{C}^{n\times n}$. There is a nonsingular matrix $S \in \mathbb{C}^{n\times n}$ such that*

$$A = SJS^{-1}$$

where J is a block diagonal matrices with Jordan blocks on the diagonal blocks (see a textbook for precise meaning). In particular, J is upper triangular with bandwidth at most 2. If $A \in \mathbb{R}^{n\times n}$ and has real eigenvalues, then S (and J) can be taken to be real valued. The Jordan matrix J is unique up to permutations of its Jordan blocks.

The Jordan blocks have the eigenvalue on the diagonal, and possibly 1's above the diagonal (this depends a bit on how you count the eigenvalues). These 1's arise from when the algebraic multiplicity of the eigenvalues (as repeated roots of the characteristic polynomial) exceed the geometric multiplicity (the dimension spanned by their eigenvectors).

Misc/summary

To summarize,

- If a matrix is **diagonalizable**, we can write it as $A = SDS^{-1}$, which is the eigenvalue decomposition. If it is not diagonalizable, we must resort to the **Jordan form** $A = SJS^{-1}$ where J is not quite diagonal.

- If a matrix is **normal**, then we can unitarily diagonalize it, $A = UDU^*$. If it is not normal, we must resort to the **Schur factorization**, $A = UTU^*$, where T is upper triangular.

Finally, we include a result that is well known in numerical analysis but not as common in more abstract linear algebra classes.

Theorem (Gershgorin's disc theorem, 6.1.1 in H&J)**.** *Let* $A \in \mathbb{C}^{n\times n}$ *with entries* $A = [a_{ij}]$*, and define the* deleted absolute row sums *of* A *as*

$$r_{\neq i} = \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|, \quad 1 \le i \le n$$

and define the Gershgorin discs as $D_i = \{z \in \mathbb{C} : |z - a_{ii}| \le r_{\neq i}\}$*. Then every eigenvalue of* A *is located in the union of all* n *discs,* $\cup_{i=1}^{n} D_i$*. Furthermore, if a union of* k *of these discs forms a connected region that is disjoint from the remaining* $n-k$ *discs, then there are precisely* k *eigenvalues of* A *in this region.*

In particular, a strictly diagonally dominant matrix is invertible, since 0 is not in any of the Gershgorin discs. Also note that since A and A^* have the same eigenvalues, the theorem applies whether you define deleted row sums or deleted column sums.