Jack Poulson
Fast Linear Algebra
Lecture 2: January 9, 2013
Review of classes of matrices

1 Important classes of matrices

1.1 Diagonalizable matrices

A square matrix A is called *diagonalizable* if it can be decomposed as

$$A = XDX^{-1},$$

where X is the matrix of *eigenvectors* of A, and D is the diagonal matrix of *eigenvalues*. It is well-known that diagonalizable matrices form a *dense* subset of the set of square matrices since all eigenvalues can be rendered unique with an arbitrarily small perturbation of the matrix (please recall that non-diagonalizable matrices must have repeated eigenvalues).

Every non-zero triangular matrix T which is zero on its main diagonal can easily be shown to be non-diagonalizable using the well-known fact that the eigenvalues of a triangular matrix lie on its diagonal:

$$T = X0X^{-1} = 0$$

is clearly a contradiction. An arbitrarily small perturbation can be applied to the diagonal of the triangular matrix T to give it small, unique eigenvalues, but the eigenvector matrix would be very close to linearly dependent.

1.2 Schur decomposition

It is also well-known that, while not all matrices are diagonalizable, the *Schur decomposition*,

$$A = QTQ^H,$$

where Q is unitary and T is upper-triangular, always exists. Recall that the eigenvalues of A lie along T's diagonal. The Schur decomposition is the preferred analogue of the eigenvalue decomposition for general square matrices, and the columns of Q are called the *Schur vectors*.

1.3 Normal matrices

It can easily be shown that the class of *normal matrices*, which satisfy

$$AA^H = A^H A,$$

is precisely the class of matrices where the Schur decomposition is also an eigenvalue decomposition (that is, T is diagonal). One direction is easy: Suppose that $A = QDQ^H$, then

$$A^H A = (QDQ^H)^H(QDQ^H) = QD^H Q^H QDQ^H = Q|D|^2 Q^H,$$

and, AA^H yields the same result.

The other direction requires slightly more care: Suppose that $AA^H = A^H A$. Then, using the Schur decomposition,

$$AA^H = (QTQ^H)(QT^H Q^H) = QTT^H Q^H$$

must be equal to

$$A^H A = (QT^H Q^H)(QTQ^H) = QT^H TQ^H,$$

which is equivalent to the requirement

$$TT^H = T^H T.$$

Partitioning T as

$$T = \begin{pmatrix} \tau_{1,1} & t_{1,2} \\ 0 & T_{2,2} \end{pmatrix},$$

the expression $TT^H = T^H T$ becomes

$$\begin{pmatrix} \tau_{1,1} & t_{1,2} \\ 0 & T_{2,2} \end{pmatrix} \begin{pmatrix} \bar{\tau}_{1,1} & 0 \\ t_{1,2}^H & T_{2,2}^H \end{pmatrix} = \begin{pmatrix} \bar{\tau}_{1,1} & 0 \\ t_{1,2}^H & T_{2,2}^H \end{pmatrix} \begin{pmatrix} \tau_{1,1} & t_{1,2} \\ 0 & T_{2,2} \end{pmatrix},$$

and equating the top-left entries of both products yields

$$|\tau_{1,1}|^2 + t_{1,2} t_{1,2}^H = |\tau_{1,1}|^2,$$

which shows that $t_{1,2} = 0$. We can recurse on

$$T_{2,2} T_{2,2}^H = T_{2,2}^H T_{2,2}$$

to show that T is in fact diagonal.

1.4 Hermitian matrices

A *Hermitian* matrix satisfies the property

$$A = A^H,$$

which clearly implies that $AA^H = A^H A$. Thus, all Hermitian matrices are *normal*, and, in fact, their eigenvalues are all real:

$$A = A^H \implies QDQ^H = Q\bar{D}Q^H \implies D = \bar{D}.$$

1.5 Unitary matrices

A matrix is *unitary* if its adjoint is its inverse, i.e.,

$$AA^H = A^H A = I.$$

Clearly all unitary matrices are then normal, and we may easily see that their eigenvalues all lie on the unit circle in the complex plane:

$$(QDQ^H)(Q\bar{D}Q^H) = I \implies Q|D|^2 Q^H = I \implies |D|^2 = I.$$

2 Our first algorithm: Cholesky factorization

The simplest of all factorization schemes is *Cholesky factorization*, which produces the decomposition

$$A = U^H U,$$

where U is upper-triangular ($A = LL^H$, where L is lower-triangular, is also standard). Since such a decomposition would imply that $(Ax, x) = (U^H U x, x) = (Ux, Ux) = \|Ux\|_2^2 \geq 0$, we know that A cannot have any negative eigenvalues (such a matrix is called *positive semi-definite*). Since $U^H U$ is clearly Hermitian, we know that A must be both Hermitian and Positive Semi-Definite (HPSD) for Cholesky factorization to be possible. It turns out that algorithms which handle the case where A is singular are significantly more complicated, and so we will assume that A is Hermitian Positive-Definite (HPD), which, if you recall, implies that every principal minor (square contiguous submatrix containing the top-left entry) is also HPD. An immediate consequence is that the top-left entry must be real and positive.

Let us expose the upper-left entry of both A and U in the equation $A = U^H U$ to find:

$$\begin{pmatrix} \alpha_{1,1} & a_{1,2} \\ a_{2,1} & A_{2,2} \end{pmatrix} = \begin{pmatrix} \bar{\upsilon}_{1,1} & 0 \\ u_{1,2}^H & U_{2,2}^H \end{pmatrix} \begin{pmatrix} \upsilon_{1,1} & u_{1,2} \\ 0 & U_{2,2} \end{pmatrix} = \begin{pmatrix} |\upsilon_{1,1}|^2 & \bar{\upsilon}_{1,1} u_{1,2} \\ \upsilon_{1,1} u_{1,2}^H & u_{1,2}^H u_{1,2} + U_{2,2}^H U_{2,2} \end{pmatrix}.$$

From the top-left equation, $\alpha_{1,1} = |\upsilon_{1,1}|^2$, we know that a valid choice of $\upsilon_{1,1}$ would be $\sqrt{\alpha_{1,1}}$, since we know that $\alpha_{1,1} > 0$. The top-right equation, $a_{1,2} = \bar{\upsilon}_{1,1} u_{1,2}$, shows that we may now compute $u_{1,2}$ as $a_{1,2}/\upsilon_{1,1}$ ($\upsilon_{1,1}$ is real by choice). Lastly, if we now have access to $u_{1,2}$, then the bottom-right equation,

$$A_{2,2} = u_{1,2}^H u_{1,2} + U_{2,2}^H U_{2,2},$$

can be reduced to

$$S_{2,2} = U_{2,2}^H U_{2,2},$$

where

$$S_{2,2} \equiv A_{2,2} - u_{1,2}^H u_{1,2} = A_{2,2} - a_{2,1} \alpha_{1,1}^{-1} a_{1,2}$$

is known as the *Schur complement* of $A_{2,2}$. Computing $U_{2,2}$ is thus a matter of computing the Cholesky factor of $S_{2,2}$, which provides us with a simple recursive algorithm:

1. $\upsilon_{1,1} := \sqrt{\alpha_{1,1}}$
2. $u_{1,2} := a_{1,2}/\upsilon_{1,1}$
3. $S_{2,2} := A_{2,2} - u_{1,2}^H u_{1,2}$
4. $U_{2,2} := \mathrm{Cholesky}(S_{2,2})$

It can easily be seen that the third step takes only k^2 operations when $A_{2,2}$ is $k \times k$, and that the overall cost of the algorithm is closely modeled by

$$\sum_{k=1}^{n-1} k^2 \approx \frac{1}{3} n^3.$$