1. Some Useful Matrix Decompositions

This short summary of orthogonal matrices, eigenvalues and singular values is restricted to square real matrices \( A \) of dimension \( N \times N \), as used in our book.

Orthogonal Matrices and Projections

A real, square matrix \( U \in \mathbb{R}^{N \times N} \) is orthogonal if its inverse equals its transpose, \( U^{-1} = U^T \). Consequently we have the two relations

\[
U^T U = I \quad \text{and} \quad U U^T = I.
\]

The columns of \( U \) are orthonormal, i.e., they are orthogonal and the 2-norm of each column is one. To see this, let \( u_i \) denote the \( i \)th column of \( U \) so that \( U = [u_1 \ u_2 \ldots \ u_N] \). Then the relation \( U^T U = I \) implies that

\[
\begin{align*}
    u_i^T u_j &= \delta_{ij} = \begin{cases} 
        1 & \text{if } i = j \\
        0 & \text{otherwise}
    \end{cases}.
\end{align*}
\]

An orthogonal matrix is perfectly well conditioned; its condition number (in any norm) is one. Moreover, any operation with its inverse merely involves a matrix product with its transpose.

An orthogonal transformation is accomplished by multiplication with an orthogonal matrix. Such a transformation leaves the 2-norm unchanged, because

\[
\|U x\|_2 = \left( (Ux)^T (Ux) \right)^{1/2} = (x^T x)^{1/2} = \|x\|_2.
\]

An orthogonal transformation can be considered as a change of basis between the “canonical” basis \( e_1, e_2, \ldots, e_N \) in \( \mathbb{R}^N \) (where \( e_i \) is the \( i \)th column of the identity matrix) and the basis \( u_1, u_2, \ldots, u_N \) given by the orthonormal columns of \( U \).

Specifically, for an arbitrary vector \( x \in \mathbb{R}^n \) we can find scalars \( z_1, \ldots, z_n \) so that

\[
x = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} = \sum_{i=1}^{N} x_i e_i = \sum_{i=1}^{N} z_i u_i = U z
\]

and it follows immediately that the coordinates \( z_i \) in the new basis are the elements of the vector

\[
z = U^T x.
\]

Because they do not distort the size of vectors, orthogonal transformations are valuable tools in numerical computations.

For any \( k \) less than \( N \) the vectors \( u_1, \ldots, u_k \) span a \( k \)-dimension subspace \( S_k \subset \mathbb{R}^N \). The orthogonal projection \( x_k \in \mathbb{R}^N \) of an arbitrary vector \( x \in \mathbb{R}^N \) onto this subspace is the unique vector in \( S_k \) which is closest to \( x \) in the 2-norm, and it is computed as

\[
x_k = U_k U_k^T x, \quad \text{with} \quad U_k = [u_1 \ u_2 \ldots \ u_k].
\]

The matrix \( U_k U_k^T \), which is \( N \times N \) and has rank \( k \), is called an orthogonal projector.
The Spectral Decomposition

A real, symmetric matrix $A = A^T$ always has an eigenvalue decomposition (or spectral decomposition) of the form

$$A = U \Lambda U^T,$$

where $U$ is orthogonal, and $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N)$ is a diagonal matrix whose diagonal elements $\lambda_i$ are the eigenvalues of $A$. A real symmetric matrix always has real eigenvalues. The columns $u_i$ of $U$ are the eigenvectors of $A$, and the eigenpairs $(\lambda_i, u_i)$ satisfy

$$A u_i = \lambda_i u_i, \quad i = 1, \ldots, N.$$

The matrix $A$ represents a linear mapping from $\mathbb{R}^N$ onto itself, and the geometric interpretation of the eigenvalue decomposition is that $U$ represents a new, orthonormal basis in which this mapping is the diagonal matrix $\Lambda$. In particular, each basis vector $u_i$ is mapped to a vector in the same direction, namely, the vector $A u_i = \lambda_i u_i$.

A real square matrix is normal if it satisfies $A A^T = A^T A$. Important examples of normal matrices are symmetric, circulant and Hankel matrices. A normal matrix has a spectral decomposition of the form

$$A = \tilde{U} \Lambda \tilde{U}^*,$$

where the complex matrix $\tilde{U}$ is unitary, i.e.,

$$\tilde{U}^{-1} = \tilde{U}^* = \text{conj}(\tilde{U})^T,$$

and $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N)$ is a diagonal matrix containing the (possibly complex) eigenvalues of $A$. (Note that orthogonal matrices are included in the set of unitary matrices.) If $A$ is real and normal, then its eigenvalues are either real or appear in complex conjugate pairs. The columns $\tilde{u}_i$ of $\tilde{U}$ are the eigenvectors of $A$. We note that a unitary matrix $\tilde{U}$ has orthonormal columns: $\tilde{u}_i^* \tilde{u}_j = \text{conj}(u_j)^T u_j = \delta_{ij}$. Also note that multiplication with $\tilde{U}$ leaves the 2-norm unchanged: $\|\tilde{U} x\|_2 = \|x\|_2$.

The Singular Value Decomposition (SVD)

A real matrix which is not normal cannot be diagonalized by an orthogonal or unitary matrix. It takes two orthogonal matrices $U$ and $V$ to diagonalize such a matrix, by means of the singular value decomposition,

$$A = U \Sigma V^T = \sum_{i=1}^N u_i \sigma_i v_i^T,$$

where $\Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_N)$ is a real diagonal matrix whose diagonal elements $\sigma_i$ are the singular values of $A$, while the singular vectors $u_i$ and $v_i$ are the columns of the orthogonal matrices $U$ and $V$. The singular values are nonnegative and are typically written in nonincreasing order:

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_N \geq 0.$$
We note that if $A$ is normal, then its singular values are equal to the absolute values of its eigenvalues.

The geometric interpretation of the SVD is that it provides two sets of orthogonal basis vectors – the columns of $U$ and $V$ – such that the mapping represented by $A$ becomes a diagonal matrix when expressed in these bases. Specifically, we have

$$A v_i = \sigma_i u_i, \quad i = 1, \ldots, N.$$  

That is, $\sigma_i$ is the “magnification” when mapping $v_i$ onto $u_i$. Any vector $x \in \mathbb{R}^N$ can be written as $x = \sum_{i=1}^{N} (v_i^T x) v_i$, and it follows that its image is given by

$$A x = \sum_{i=1}^{N} (v_i^T x) A v_i = \sum_{i=1}^{N} \sigma_i (v_i^T x) u_i.$$  

If $A$ has an inverse, then the mapping of the inverse also defines a diagonal matrix:

$$A^{-1} u_i = \sigma_i^{-1} v_i,$$

so that $\sigma_i^{-1}$ is the “magnification” when mapping $u_i$ back onto $v_i$. Similarly, any vector $b \in \mathbb{R}^N$ can be written as $x = \sum_{i=1}^{N} (u_i^T b) u_i$, and it follows that the vector $A^{-1} b$ is given by

$$A^{-1} b = \sum_{i=1}^{N} (u_i^T b) A^{-1} u_i = \sum_{i=1}^{N} \frac{u_i^T b}{\sigma_i} v_i.$$  

Similar relations can easily be derived for the spectral decompositions.

**Rank, Conditioning, and Truncated SVD**

The rank of a matrix is equal to the number of nonzero singular values: $r = \text{rank}(A)$ means that

$$\sigma_r > 0, \quad \sigma_{r+1} = 0.$$  

The matrix $A$ has full rank (and, therefore, an inverse) only if all of its singular values are nonzero. If $A$ is rank deficient then the system $A x = b$ may not be compatible; in other words, there may be no vector $x$ that solves the problem. The columns of $U_r = [u_1 \ u_2 \ \ldots \ u_r]$ form an orthonormal basis for the range of $A$, and the system $A x = b_r$ with $b_r = U_r^T b$ is the closest compatible system. This compatible system has infinitely many solutions, and the solution of minimum 2-norm is

$$x_r = \sum_{i=1}^{r} \frac{u_i^T b}{\sigma_i} v_i.$$  

Consider now a perturbed version $A \tilde{x} = \tilde{b}$ of the original system $A x = b$, in which the perturbed right-hand side is given by $\tilde{b} = b + e$. If $A$ has full rank then the perturbed solution is given by $\tilde{x} = A^{-1} \tilde{b} = x + A^{-1} e$, and we need an upper bound for the relative perturbation $\|x - \tilde{x}\|_2/\|x\|_2$. The worst-case situation arises
when \( \mathbf{b} \) is in the direction of the left singular vector \( \mathbf{u}_1 \) while the perturbation \( \mathbf{e} \) is solely in the direction of \( \mathbf{u}_N \), and it follows that the perturbation bound is given by

\[
\frac{\| \mathbf{x} - \tilde{\mathbf{x}} \|_2}{\| \mathbf{x} \|_2} \leq \text{cond}(\mathbf{A}) \frac{\| \mathbf{e} \|_2}{\| \mathbf{b} \|_2}, \quad \text{where} \quad \text{cond}(\mathbf{A}) = \frac{\sigma_1}{\sigma_N}.
\]

The quantity \( \text{cond}(\mathbf{A}) \) is the condition number of \( \mathbf{A} \). The larger the condition number, the more sensitive the system is to perturbations of the right-hand side.

The smallest singular value \( \sigma_N \) measures how “close” \( \mathbf{A} \) is to a singular matrix (and \( \sigma_N = 0 \) when \( \mathbf{A} \) is singular). A perturbation of \( \mathbf{A} \) with a matrix \( \mathbf{E} \), whose elements are of the order \( \sigma_N \), can make \( \mathbf{A} \) rank deficient. The existence of one or more small singular values (small compared to the largest singular value \( \sigma_1 \)) therefore indicates that \( \mathbf{A} \) is “almost” singular.

In this case, it is often recommended to replace the ill-conditioned matrix \( \mathbf{A} \) with a nearby but exactly rank-deficient matrix \( \mathbf{A}_k \) whose rank \( k \) cannot be reduced by small perturbations. The typical choice of \( \mathbf{A}_k \) is the truncated SVD (TSVD) matrix

\[
\mathbf{A}_k = \mathbf{U}_k \mathbf{U}_k^T \mathbf{A} = \sum_{i=1}^{k} \mathbf{u}_i \sigma_i \mathbf{v}_i^T.
\]

The rank \( k \) of \( \mathbf{A}_k \) is chosen such that \( \sigma_k \) – which measures how “close” \( \mathbf{A}_k \) is to a singular matrix – is larger than the perturbations (the errors) in the original matrix \( \mathbf{A} \). The minimum-norm solution to the corresponding compatible system \( \mathbf{A}_k \mathbf{x} = \mathbf{b}_k \) with \( \mathbf{b}_k = \mathbf{U}_k \mathbf{U}_k^T \mathbf{b} \) is called the TSVD solution, and it is given by

\[
\mathbf{x}_k = \mathbf{U}_k \mathbf{U}_k^T \mathbf{x} = \sum_{i=1}^{k} \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{v}_i.
\]
2. The DFT and Smoothing Norms

The following is a derivation of equation (7.9) in Section 7.3 for the efficient computation of the Tikhonov solution with a smoothing norm \( \|Dx\|_2 \) that involves partial derivatives. We consider the case of periodic boundary conditions where the DFT matrix \( F = F_r \otimes F_c \) diagonalizes the matrix \( A \), i.e.,

\[
A = F^* \Lambda_A F,
\]

in which the diagonal matrix \( \Lambda_A \) contains the eigenvalues \( \lambda_i \) of \( A \). These eigenvalues are computed as described in section 4.2.

For periodic boundary conditions, the one-dimensional DFT matrix \( F_c \) diagonalizes the first and second derivative matrices \( D_{1,m} \) (7.7) and \( D_{2,m} \) (7.6), i.e.,

\[
D_{q,m} = F_r^* A_{D_{q,m}} F_c, \quad q = 1, 2,
\]

where the diagonal matrix \( A_{D_{q,m}} \) contains the eigenvalues \( \lambda_{q,k} \) of \( D_{q,m} \). These eigenvalues can be pre-computed by the following expressions

\[
\lambda_{q,k} = \begin{cases} 
\exp(2k\pi i/m) - 1, & q = 1 \\
2 \cos(2k\pi/m) - 2, & q = 2 
\end{cases} \quad \text{for} \quad k = 1, \ldots, m
\]

in which \( i = \sqrt{-1} \) denotes the imaginary unit. It follows that several choices of the matrix \( D \) have simple expressions in terms of Kronecker products; for example

\[
I_n \otimes D_{q,m} = (F_r^* F_r) \otimes (F_c^* A_{D_{q,m}} F_c) = (F_r^* \otimes F_c^*) (I_n \otimes A_{D_{q,m}}) (F_r \otimes F_c) = F^* (I_n \otimes A_{D_{q,m}}) F.
\]

Similarly, we obtain

\[
D_{q,n} \otimes I_m = F^* (A_{D_{q,n}} \otimes I_m) F
\]

\[
I_n \otimes D_{2,m} + D_{2,n} \otimes I_m = F^* (I_n \otimes A_{D_{2,m}} + A_{D_{2,n}} \otimes I_m) F
\]

\[
\begin{bmatrix} I_n \otimes D_{q,m} \\ D_{q,n} \otimes I_m \end{bmatrix} = \begin{bmatrix} F^* & 0 \\ 0 & F^* \end{bmatrix} \begin{bmatrix} I_n \otimes A_{D_{q,m}} \\ A_{D_{q,n}} \otimes I_m \end{bmatrix} F.
\]

We can use the above relations to derive a simple expression for the Tikhonov solution to (7.3). We need the following result:

\[
A^T = A^*
\]

\[
= F^* \text{conj}(A_A) F
\]

\[
= F^* \text{conj}(A_A) A_A^{-1} F
\]

\[
= F^* |A_A|^{-2} A_A^{-1} F
\]

where \( |A_A|^2 \) denotes a diagonal matrix whose elements are \( |\lambda_i|^2 \). It follows immediately that

\[
A^T A = F^* \text{conj}(A_A) A_A F = F^* |A_A|^2 F.
\]
A similar result holds for the matrix $D$, depending on its form. For example, if $D = I_n \otimes D_{q,m} = F^* (I_n \otimes A_{D_{q,m}}) F$ then

$$D^T D = D^* = F^* (I_n \otimes \text{conj}(A_{D_{q,m}})) F$$

and hence

$$D^T D = F^* (I_n \otimes \text{conj}(A_{D_{q,m}})) (I_n \otimes A_{D_{q,m}}) F = F^* (I_n \otimes \text{conj}(A_{D_{q,m}}) A_{D_{q,m}}) F = F^* (I_n \otimes |A_{D_{q,m}}|^2) F.$$ 

Putting the above relations together, we arrive at the following expression for the Tikhonov solution

$$x_{\alpha,D} = (A^T A + \alpha^2 D^T D)^{-1} A^T b = F^* \left( |A_A|^2 \left( |A_A|^2 + \alpha^2 (I_n \otimes |A_{D_{q,m}}|^2) \right)^{-1} \right) A_A^{-1} F b.$$ 

There are similar expressions for the other choices of the matrix $D$. If $D = D_{q,n} \otimes I_m = F^* (A_{D_{q,n}} \otimes I_m) F$ then it follows immediately that

$$D^T D = F^* (|A_{D_{q,n}}|^2 \otimes I_m) F,$$

and if we use a sum of squared norm, represented by

$$D = \begin{bmatrix} I_n \otimes D_{q,m} \\ D_{q,n} \otimes I_m \end{bmatrix} = \begin{bmatrix} F^* & 0 \\ 0 & F^* \end{bmatrix} \begin{bmatrix} I_n \otimes A_{D_{q,m}} \\ A_{D_{q,n}} \otimes I_m \end{bmatrix} F,$$

then we obtain

$$D^T D = F^* \left( I_n \otimes |A_{D_{q,m}}|^2 + |A_{D_{q,n}}|^2 \otimes I_m \right) F.$$ 

Finally if $D = I_n \otimes D_{2,m} + D_{2,n} \otimes I_m = F^* (I_n \otimes A_{D_{2,m}} + A_{D_{2,n}} \otimes I_m) F$ (approximating the Laplacian) then we obtain

$$D^T D = F^* \left( I_n \otimes A_{D_{2,m}}^2 + A_{D_{2,n}}^2 \otimes I_m + 2 A_{D_{2,n}} \otimes A_{D_{2,m}} \right) F.$$ 

The absolute value is not necessary here because the eigenvalues are real.

We can summarize these results in the expression (7.9) for the Tikhonov solution

$$x_{\alpha,D} = F^* \left( |A_A|^2 \left( |A_A|^2 + \alpha^2 \Delta \right)^{-1} \right) A_A^{-1} F b,$$

where the diagonal matrix $\Delta$ takes one of the four forms shown in the last column of Table 7.2. Note that the filtering matrix $|A_A|^2 \left( |A_A|^2 + \alpha^2 \Delta \right)^{-1}$ is diagonal.