The SVD of *A*

As stated in Theorem 5.4, the SVD of the $m \times n$ matrix A ($m \geq n$), $A = U\Sigma V^*$, is related to the eigenvalue decomposition of the matrix A^*A,

$$A^*A = V\Sigma^*\Sigma V^*.$$

Thus, mathematically speaking, we might calculate the SVD of A as follows:

1. Form A^*A;
2. Compute the eigenvalue decomposition $A^*A = V\Lambda V^*$;
3. Let Σ be the $m \times n$ nonnegative diagonal square root of Λ;
4. Solve the system $U\Sigma = AV$ for unitary U (e.g., via QR factorization).

- This reduction of the SVD problem to a eigenvalue problem is unstable against small perturbations. (Explanation is omitted and details can be seen from “the Book”).

An Alternative Reduction

There is an alternative, stable way to reduce the SVD to an eigenvalue problem. Assume that A is square, with $m = n$; this is no essential restriction, since we shall see that rectangular singular value problems can be reduced to square ones. Consider the $2m \times 2m$ hermitian matrix

$$H = \begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix} \tag{31.2}$$

mentioned earlier in Exercise 5.4. Since $A = U\Sigma V^*$ implies $AV = U\Sigma$ and $A^*U = V\Sigma^* = V\Sigma$, we have the block 2×2 equation

$$\begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix} \begin{bmatrix} V & V \\ U & -U \end{bmatrix} = \begin{bmatrix} V & V \\ U & -U \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{bmatrix}, \tag{31.3}$$

which amounts to an eigenvalue decomposition of H. Thus we see that the singular values of A are the absolute values of the eigenvalues of H, and the singular vectors of A can be extracted from the eigenvectors of H.

Two-Phase Approach

We have seen that hermitian eigenvalue problems are usually solved by a two-phase computation: first reduce the matrix to tridiagonal form, then diagonalize the tridiagonal matrix. Since the work of Golub, Kahan, and others in the 1960s, an analogous two-phase approach has been standard for the SVD. The matrix is brought into bidiagonal form, and then the bidiagonal matrix is diagonalized:

$$\underset{A}{\begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix}} \xrightarrow{\text{Phase 1}} \underset{B}{\begin{bmatrix} \times & \times & & \\ & \times & \times & \\ & & \times & \times \\ & & & \times \\ & & & \\ & & & \end{bmatrix}} \xrightarrow{\text{Phase 2}} \underset{\Sigma}{\begin{bmatrix} \times & & & \\ & \times & & \\ & & \times & \\ & & & \times \\ & & & \\ & & & \end{bmatrix}}.$$

Phase 1: Golub-Kahan Bidiagonalization

In Phase 1 of the SVD computation, we bring A into bidiagonal form by applying distinct unitary operations on the left and right. Note how this differs from the computation of eigenvalues, where the same unitary operations must be applied on both sides so that each step is a similarity transformation. In that case, it was only possible to introduce zeros below the first subdiagonal. Here, we are able to completely triangularize and also introduce zeros above the first superdiagonal.

Phase 1: Golub-Kahan Bidiagonalization

The simplest method for accomplishing this, *Golub–Kahan bidiagonalization*, proceeds as follows. Householder reflectors are applied alternately on the left and the right. Each left reflection introduces a column of zeros below the diagonal. The right reflection that follows introduces a row of zeros to the right of the first superdiagonal, leaving intact the zeros just introduced in the column. For example, for a 6×4 matrix, the first two pairs of reflections look like this:

$$
\underset{A}{\begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix}}
\xrightarrow{U_1^* \cdot}
\underset{U_1^* A}{\begin{bmatrix} \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \end{bmatrix}}
\xrightarrow{\cdot V_1}
\underset{U_1^* A V_1}{\begin{bmatrix} \times & \boldsymbol{\times} & \mathbf{0} & \mathbf{0} \\ & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \end{bmatrix}}
$$

$$
\xrightarrow{U_2^* \cdot}
\underset{U_2^* U_1^* A V_1}{\begin{bmatrix} \times & \times & & \\ & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ & \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} \\ & \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} \\ & \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} \\ & \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} \end{bmatrix}}
\xrightarrow{\cdot V_2}
\underset{U_2^* U_1^* A V_1 V_2}{\begin{bmatrix} \times & \times & & \\ & \times & \boldsymbol{\times} & \mathbf{0} \\ & & \boldsymbol{\times} & \boldsymbol{\times} \\ & & \boldsymbol{\times} & \boldsymbol{\times} \\ & & \boldsymbol{\times} & \boldsymbol{\times} \\ & & \boldsymbol{\times} & \boldsymbol{\times} \end{bmatrix}}.
$$

Phase 1: Golub-Kahan Bidiagonalization

At the end of this process, n reflectors have been applied on the left and $n-2$ on the right. The pattern of floating point operations resembles two Householder QR factorizations interleaved with each other, one operating on the $m \times n$ matrix A, the other on the $n \times m$ matrix A^*. The total operation count is therefore twice that of a QR factorization (10.9), i.e.,

$$\text{Work for Golub–Kahan bidiagonalization:} \quad \sim 4mn^2 - \frac{4}{3}n^3 \text{ flops.} \qquad (31.4)$$

- There are faster methods.

Phase 2

In Phase 2 of the computation of the SVD, the SVD of the bidiagonal matrix B is determined. From the 1960s to the 1990s, the standard algorithm for this was a variant of the QR algorithm. More recently, divide-and-conquer algorithms have also become competitive, and in the future, they are likely to become the standard. We shall not give details.