Chapter 6
Singular Value Decomposition

Even if the QR decomposition is very useful for solving least squares problems and has excellent stability properties, it has the drawback that it treats the rows and columns of the matrix differently: it gives a basis only for the *column space.* The singular value decomposition (SVD) deals with the rows and columns in a symmetric fashion, and therefore it supplies more information about the matrix. It also "orders" the information contained in the matrix so that, loosely speaking, the "dominating part" becomes visible. This is the property that makes the SVD so useful in data mining and many other areas.

6.1 The Decomposition

Theorem 6.1 (SVD). *Any $m \times n$ matrix A, with $m \geq n$, can be factorized*

$$A = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^T, \tag{6.1}$$

where $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ are orthogonal, and $\Sigma \in \mathbb{R}^{n \times n}$ is diagonal,

$$\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n),$$

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0.$$

Proof. The assumption $m \geq n$ is no restriction: in the other case, just apply the theorem to A^T. We give a proof along the lines of that in [42]. Consider the maximization problem

$$\sup_{\| x \|_2 = 1} \| Ax \|_2.$$

Since we are seeking the supremum of a continuous function over a closed set, the supremum is attained for some vector x. Put $Ax = \sigma_1 y$, where $\| y \|_2 = 1$ and

$\sigma_1 = \| A \|_2$ (by definition). Using Proposition 4.7 we can construct orthogonal matrices

$$Z_1 = (y\, \bar{Z}_2) \in \mathbb{R}^{m\times m}, \quad W_1 = (x\, \bar{W}_2) \in \mathbb{R}^{n\times n}.$$

Then

$$Z_1^T A W_1 = \begin{pmatrix} \sigma_1 & y^T A \bar{W}_2 \\ 0 & \bar{Z}_2^T A \bar{W}_2 \end{pmatrix},$$

since $y^T A x = \sigma_1$, and $Z_2^T A x = \sigma_1 \bar{Z}_2^T y = 0$. Put

$$A_1 = Z_1^T A W_1 = \begin{pmatrix} \sigma_1 & w^T \\ 0 & B \end{pmatrix}.$$

Then

$$\frac{1}{\sigma_1^2 + w^T w} \left\| A_1 \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2^2 = \frac{1}{\sigma_1^2 + w^T w} \left\| \begin{pmatrix} \sigma_1^2 + w^T w \\ Bw \end{pmatrix} \right\|_2^2 \geq \sigma_1^2 + w^T w.$$

But $\| A_1 \|_2^2 = \| Z_1^T A W_1 \|_2^2 = \sigma_1^2$; therefore $w = 0$ must hold. Thus we have taken one step toward a diagonalization of A. The proof is now completed by induction. Assume that

$$B = Z_2 \begin{pmatrix} \Sigma_2 \\ 0 \end{pmatrix} W_2, \quad \Sigma_2 = \mathrm{diag}(\sigma_2, \ldots, \sigma_n).$$

Then we have

$$A = Z_1 \begin{pmatrix} \sigma_1 & 0 \\ 0 & B \end{pmatrix} W_1^T = Z_1 \begin{pmatrix} 1 & 0 \\ 0 & Z_2 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \Sigma_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & W_2^T \end{pmatrix} W_1^T.$$

Thus, by defining

$$U = Z_1 \begin{pmatrix} 1 & 0 \\ 0 & Z_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix}, \quad V = W_1 \begin{pmatrix} 1 & 0 \\ 0 & W_2 \end{pmatrix},$$

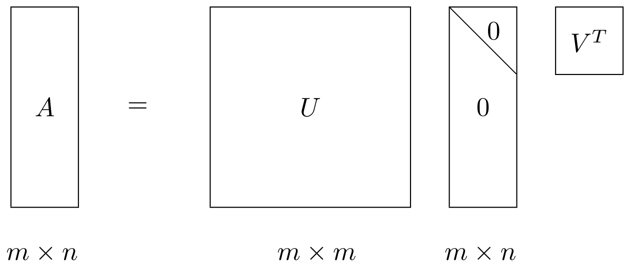
the theorem is proved. □

The columns of U and V are called *singular vectors* and the diagonal elements σ_i *singular values*.

We emphasize at this point that not only is this an important theoretical result, but also there are very efficient and accurate algorithms for computing the SVD; see Section 6.8.

The SVD appears in other scientific areas under different names. In statistics and data analysis, the singular vectors are closely related to *principal components* (see Section 6.4), and in image processing the SVD goes by the name *Karhunen–Loewe expansion*.

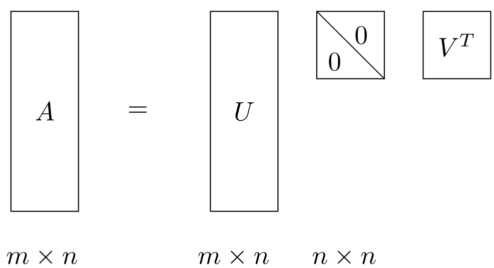
We illustrate the SVD symbolically:



With the partitioning $U = (U_1\, U_2)$, where $U_1 \in \mathbb{R}^{m\times n}$, we get the *thin SVD*,

$$A = U_1 \Sigma V^T,$$

illustrated symbolically,



If we write out the matrix equations

$$AV = U_1\Sigma, \qquad A^T U_1 = V\Sigma$$

column by column, we get the equivalent equations

$$Av_i = \sigma_i u_i, \qquad A^T u_i = \sigma_i v_i, \qquad i = 1, 2, \dots, n.$$

The SVD can also be written as an expansion of the matrix:

$$A = \sum_{i=1}^{n} \sigma_i u_i v_i^T. \tag{6.2}$$

This is usually called the *outer product form*, and it is derived by starting from the thin version:

$$A = U_1 \Sigma V^T = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix} \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n^T \end{pmatrix} \begin{pmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{pmatrix}$$

$$= \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix} \begin{pmatrix} \sigma_1 v_1^T \\ \sigma_2 v_2^T \\ \vdots \\ \sigma_n v_n^T \end{pmatrix} = \sum_{i=1}^{n} \sigma_i u_i v_i^T.$$

The outer product form of the SVD is illustrated as

$$A = \sum_{i=1}^{n} \sigma_i u_i v_i^T = \qquad + \qquad + \quad \cdots .$$

Example 6.2. We compute the SVD of a matrix with full column rank:

```
A =  1     1
     1     2
     1     3
     1     4
 >> [U,S,V]=svd(A)

U = 0.2195   -0.8073    0.0236    0.5472
    0.3833   -0.3912   -0.4393   -0.7120
    0.5472    0.0249    0.8079   -0.2176
    0.7110    0.4410   -0.3921    0.3824

S = 5.7794         0
         0    0.7738
         0         0
         0         0

V = 0.3220   -0.9467
    0.9467    0.3220
```

The thin version of the SVD is

```
>> [U,S,V]=svd(A,0)

U = 0.2195   -0.8073
    0.3833   -0.3912
    0.5472    0.0249
    0.7110    0.4410

S = 5.7794         0
         0    0.7738

V = 0.3220   -0.9467
    0.9467    0.3220
```
■

The matrix 2-norm was defined in Section 2.4. From the proof of Theorem 6.1 we know already that $\| A \|_2 = \sigma_1$. This is such an important fact that it is worth a separate proposition.

Proposition 6.3. *The 2-norm of a matrix is given by*

$$\| A \|_2 = \sigma_1.$$

Proof. The following is an alternative proof. Without loss of generality, assume that $A \in \mathbb{R}^{m\times n}$ with $m \geq n$, and let the SVD of A be $A = U\Sigma V^T$. The norm is invariant under orthogonal transformations, and therefore

$$\| A \|_2 = \| \Sigma \|_2.$$

The result now follows, since the 2-norm of a diagonal matrix is equal to the absolute value of the largest diagonal element:

$$\| \Sigma \|_2^2 = \sup_{\| y \|_2=1} \| \Sigma y \|_2^2 = \sup_{\| y \|_2=1} \sum_{i=1}^{n} \sigma_i^2 y_i^2 \leq \sigma_1^2 \sum_{i=1}^{n} y_i^2 = \sigma_1^2$$

with equality for $y = e_1$. ∎

6.2 Fundamental Subspaces

The SVD gives orthogonal bases of the four fundamental subspaces of a matrix. The *range of the matrix* A is the linear subspace

$$\mathcal{R}(A) = \{y \mid y = Ax, \text{ for arbitrary } x\}.$$

Assume that A has rank r:

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$

Then, using the outer product form, we have

$$y = Ax = \sum_{i=1}^{r} \sigma_i u_i v_i^T x = \sum_{i=1}^{r} (\sigma_i v_i^T x) u_i = \sum_{i=1}^{r} \alpha_i u_i.$$

The *null-space of the matrix* A is the linear subspace

$$\mathcal{N}(A) = \{x \mid Ax = 0\}.$$

Since $Ax = \sum_{i=1}^{r} \sigma_i u_i v_i^T x$, we see that any vector $z = \sum_{i=r+1}^{n} \beta_i v_i$ is in the null-space:

$$Az = \left(\sum_{i=1}^{r} \sigma_i u_i v_i^T \right) \left(\sum_{i=r+1}^{n} \beta_i v_i \right) = 0.$$

After a similar demonstration for A^T we have the following theorem.

Theorem 6.4 (fundamental subspaces).

1. *The singular vectors* $u_1, u_2, \dots, u_r$ *are an orthonormal basis in* $\mathcal{R}(A)$ *and*
$$\mathrm{rank}(A) = \dim(\mathcal{R}(A)) = r.$$
2. *The singular vectors* $v_{r+1}, v_{r+2}, \dots, v_n$ *are an orthonormal basis in* $\mathcal{N}(A)$ *and*
$$\dim(\mathcal{N}(A)) = n - r.$$
3. *The singular vectors* $v_1, v_2, \dots, v_r$ *are an orthonormal basis in* $\mathcal{R}(A^T)$.
4. *The singular vectors* $u_{r+1}, u_{r+2}, \dots, u_m$ *are an orthonormal basis in* $\mathcal{N}(A^T)$.

Example 6.5. We create a rank deficient matrix by constructing a third column in the previous example as a linear combination of columns 1 and 2:

```
>> A(:,3)=A(:,1)+0.5*A(:,2)

A = 1.0000    1.0000    1.5000
    1.0000    2.0000    2.0000
    1.0000    3.0000    2.5000
    1.0000    4.0000    3.0000

>> [U,S,V]=svd(A,0)

U = 0.2612   -0.7948   -0.5000
    0.4032   -0.3708    0.8333
    0.5451    0.0533   -0.1667
    0.6871    0.4774   -0.1667

S = 7.3944         0         0
         0    0.9072         0
         0         0         0

V = 0.2565   -0.6998    0.6667
    0.7372    0.5877    0.3333
    0.6251   -0.4060   -0.6667
```

The third singular value is equal to zero and the matrix is rank deficient. Obviously, the third column of `V` is a basis vector in $\mathcal{N}(A)$:

```
>> A*V(:,3)

ans =
    1.0e-15 *
         0
   -0.2220
   -0.2220
         0
```
∎

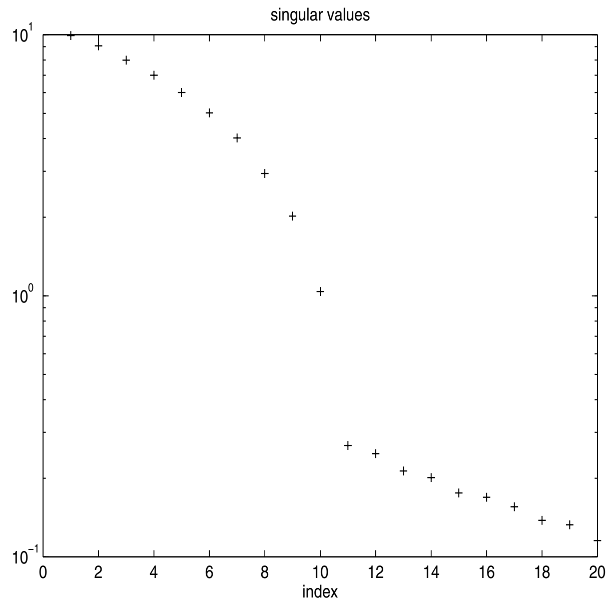


Figure 6.1. *Singular values of a matrix of rank* 10 *plus noise.*

6.3 Matrix Approximation

Assume that A is a low-rank matrix plus noise: $A = A_0 + N$, where the noise N is small compared with A_0. Then typically the singular values of A have the behavior illustrated in Figure 6.1. In such a situation, if the noise is sufficiently small in magnitude, the number of large singular values is often referred to as the *numerical rank* of the matrix. If we know the correct rank of A_0, or can estimate it, e.g., by inspecting the singular values, then we can "remove the noise" by approximating A by a matrix of the correct rank. The obvious way to do this is simply to truncate the singular value expansion (6.2). Assume that the numerical rank is equal to k. Then we approximate

$$A = \sum_{i=1}^{n} \sigma_i u_i v_i^T \approx \sum_{i=1}^{k} \sigma_i u_i v_i^T =: A_k.$$

The truncated SVD is very important, not only for removing noise but also for compressing data (see Chapter 11) and for stabilizing the solution of problems that are extremely ill-conditioned.

It turns out that the truncated SVD is the solution of approximation problems where one wants to approximate a given matrix by one of lower rank. We will consider low-rank approximation of a matrix A in two norms. First we give the theorem for the matrix 2-norm.

Theorem 6.6. *Assume that the matrix $A \in \mathbb{R}^{m \times n}$ has rank $r > k$. The matrix* *approximation problem*

$$\min_{\operatorname{rank}(Z)=k} \|A - Z\|_2$$

has the solution

$$Z = A_k := U_k \Sigma_k V_k^T,$$

where $U_k = (u_1, \ldots, u_k)$, $V_k = (v_1, \ldots, v_k)$, *and* $\Sigma_k = \operatorname{diag}(\sigma_1, \ldots, \sigma_k)$. *The minimum is*

$$\|A - A_k\|_2 = \sigma_{k+1}.$$

A proof of this theorem can be found, e.g., in [42, Section 2.5.5]. Next recall the definition of the *Frobenius matrix norm* (2.7)

$$\|A\|_F = \sqrt{\sum_{i,j} a_{ij}^2}.$$

It turns out that the approximation result is the same for this case.

Theorem 6.7. *Assume that the matrix* $A \in \mathbb{R}^{m \times n}$ *has rank* $r > k$. *The Frobenius norm matrix approximation problem*

$$\min_{\operatorname{rank}(Z)=k} \|A - Z\|_F$$

has the solution

$$Z = A_k = U_k \Sigma_k V_k^T,$$

where $U_k = (u_1, \ldots, u_k)$, $V_k = (v_1, \ldots, v_k)$, *and* $\Sigma_k = \operatorname{diag}(\sigma_1, \ldots, \sigma_k)$. *The minimum is*

$$\|A - A_k\|_F = \left(\sum_{i=k+1}^{p} \sigma_i^2 \right)^{1/2},$$

where $p = \min(m, n)$.

For the proof of this theorem we need a lemma.

Lemma 6.8. *Consider the mn-dimensional vector space* $\mathbb{R}^{m \times n}$ *with inner product*

$$\langle A, B \rangle = \operatorname{tr}(A^T B) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ij} \tag{6.3}$$

and norm

$$\|A\|_F = \langle A, A \rangle^{1/2}.$$

Let $A \in \mathbb{R}^{m\times n}$ with SVD $A = U\Sigma V^T$. Then the matrices

$$u_i v_j^T, \qquad i = 1, 2, \ldots, m, \qquad j = 1, 2, \ldots, n, \tag{6.4}$$

are an orthonormal basis in $\mathbb{R}^{m\times n}$.

Proof. Using the identities $\langle A, B\rangle = \operatorname{tr}(A^T B) = \operatorname{tr}(BA^T)$ we get

$$\langle u_i v_j^T, u_k v_l^T\rangle = \operatorname{tr}(v_j u_i^T u_k v_l^T) = \operatorname{tr}(v_l^T v_j\, u_i^T u_k) = (v_l^T v_j)\,(u_i^T u_k),$$

which shows that the matrices are orthonormal. Since there are mn such matrices, they constitute a basis in $\mathbb{R}^{m\times n}$. □

Proof (Theorem 6.7). This proof is based on that in [41]. Write the matrix $Z \in \mathbb{R}^{m\times n}$ in terms of the basis (6.4),

$$Z = \sum_{i,j} \zeta_{ij} u_i v_j^T,$$

where the coefficients are to be chosen. For the purpose of this proof we denote the elements of Σ by σ_{ij}. Due to the orthonormality of the basis, we have

$$\|A - Z\|_F^2 = \sum_{i,j} (\sigma_{ij} - \zeta_{ij})^2 = \sum_i (\sigma_{ii} - \zeta_{ii})^2 + \sum_{i\neq j} \zeta_{ij}^2.$$

Obviously, we can choose the second term as equal to zero. We then have the following expression for Z:

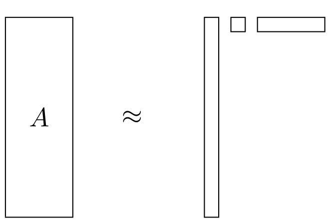
$$Z = \sum_i \zeta_{ii} u_i v_i^T.$$

Since the rank of Z is equal to the number of terms in this sum, we see that the constraint $\operatorname{rank}(Z) = k$ implies that we should have exactly k nonzero terms in the sum. To minimize the objective function, we then choose

$$\zeta_{ii} = \sigma_{ii}, \quad i = 1, 2, \ldots, k,$$

which gives the desired result. □

The low-rank approximation of a matrix is illustrated as



$= U_k \Sigma_k V_k^T.$

6.4 Principal Component Analysis

The approximation properties of the SVD can be used to elucidate the equivalence between the SVD and *principal component analysis* (*PCA*). Assume that $X \in \mathbb{R}^{m\times n}$ is a data matrix, where each column is an observation of a real-valued random vector with mean zero. The matrix is assumed to be centered, i.e., the mean of each column is equal to zero. Let the SVD of X be $X = U\Sigma V^T$. The right singular vectors v_i are called *principal components directions* of X [47, p. 62]. The vector

$$z_1 = Xv_1 = \sigma_1 u_1$$

has the largest sample variance among all normalized linear combinations of the columns of X:

$$\mathrm{Var}(z_1) = \mathrm{Var}(Xv_1) = \frac{\sigma_1^2}{m}.$$

Finding the vector of maximal variance is equivalent, using linear algebra terminology, to maximizing the Rayleigh quotient:

$$\sigma_1^2 = \max_{v\neq 0} \frac{v^T X^T X v}{v^T v}, \qquad v_1 = \arg\max_{v\neq 0} \frac{v^T X^T X v}{v^T v}.$$

The normalized variable $u_1 = (1/\sigma_1)Xv_1$ is called the *normalized first principal component* of X.

Having determined the vector of largest sample variance, we usually want to go on and find the vector of second largest sample variance *that is orthogonal to the first*. This is done by computing the vector of largest sample variance of the *deflated data matrix* $X - \sigma_1 u_1 v_1^T$. Continuing this process we can determine all the principal components in order, i.e., we compute the singular vectors. In the general step of the procedure, the subsequent principal component is defined as the vector of maximal variance subject to the constraint that it is orthogonal to the previous ones.

Example 6.9. PCA is illustrated in Figure 6.2. Five hundred data points from a correlated normal distribution were generated and collected in a data matrix $X \in \mathbb{R}^{3\times 500}$. The data points and the principal components are illustrated in the top plot of the figure. We then deflated the data matrix, $X_1 := X - \sigma_1 u_1 v_1^T$; the data points corresponding to X_1 are given in the bottom plot. ■

6.5 Solving Least Squares Problems

The least squares problem can be solved using the SVD. Assume that we have an overdetermined system $Ax \sim b$, where the matrix A has full column rank. Write the SVD

$$A = (U_1 \; U_2) \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^T,$$

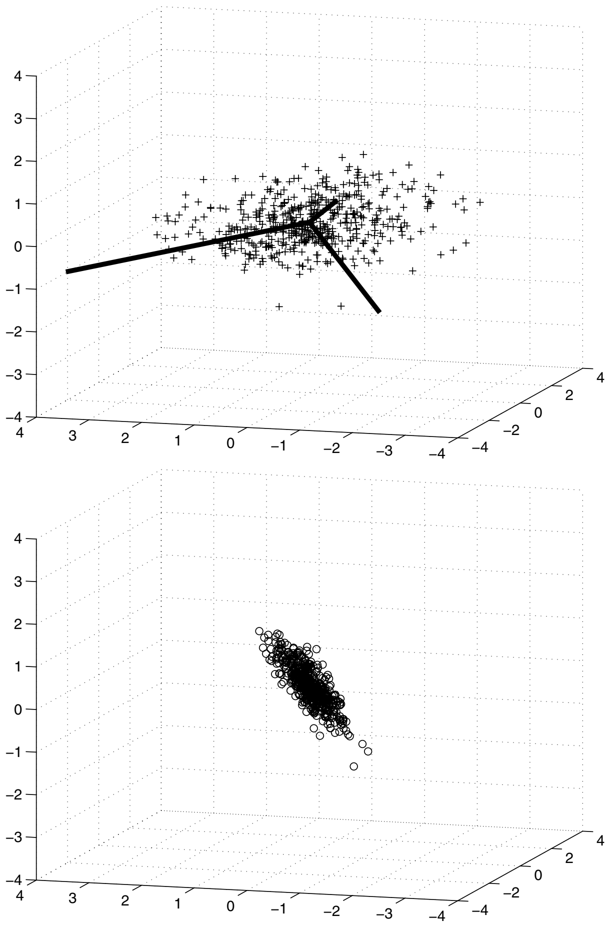


Figure 6.2. *Top: Cluster of points in $\mathbb{R}^3$ with (scaled) principal components. Bottom: same data with the contributions along the first principal component deflated.*

where $U_1 \in \mathbb{R}^{m\times n}$. Using the SVD and the fact that the norm is invariant under orthogonal transformations, we have

$$\|r\|^2 = \|b - Ax\|^2 = \left\| b - U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^T x \right\|^2 = \left\| \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} - \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} y \right\|^2,$$

where $b_i = U_i^T b$ and $y = V^T x$. Thus

$$\|r\|^2 = \|b_1 - \Sigma y\|^2 + \|b_2\|^2.$$

We can now minimize $\|r\|^2$ by putting $y = \Sigma^{-1}b_1$. The least squares solution is given by

$$x = Vy = V\Sigma^{-1}b_1 = V\Sigma^{-1}U_1^T b. \tag{6.5}$$

Recall that Σ is diagonal,

$$\Sigma^{-1} = \operatorname{diag}\left(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \dots, \frac{1}{\sigma_n}\right),$$

so the solution can also be written

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

The assumption that A has full column rank implies that all the singular values are nonzero: $\sigma_i > 0$, $i = 1, 2, \dots, n$. We also see that in this case, the solution is unique.

Theorem 6.10 (least squares solution by SVD). *Let the matrix $A \in \mathbb{R}^{m\times n}$ have full column rank and thin SVD $A = U_1\Sigma V^T$. Then the least squares problem $\min_x \| Ax - b \|_2$ has the unique solution*

$$x = V\Sigma^{-1}U_1^T b = \sum_{i=1}^{n} \frac{u_i^T b}{\sigma_i} v_i.$$

Example 6.11. As an example, we solve the least squares problem given at the beginning of Chapter 3.6. The matrix and right-hand side are

```
A = 1     1              b = 7.9700
    1     2                 10.2000
    1     3                 14.2000
    1     4                 16.0000
    1     5                 21.2000

>> [U1,S,V]=svd(A,0)

U1 =0.1600   -0.7579
    0.2853   -0.4675
    0.4106   -0.1772
    0.5359    0.1131
    0.6612    0.4035

 S = 7.6912         0
         0    0.9194

 V = 0.2669   -0.9637
    0.9637    0.2669
```

The two column vectors in A are linearly independent since the singular values are both nonzero. The least squares problem is solved using (6.5):

```
>> x=V*(S\(U1'*b))

x = 4.2360
    3.2260
```

■

6.6 Condition Number and Perturbation Theory for the Least Squares Problem

The condition number of a rectangular matrix is defined in terms of the SVD. Let A have rank r, i.e., its singular values satisfy

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_p = 0,$$

where $p = \min(m, n)$. Then the condition number is defined

$$\kappa(A) = \frac{\sigma_1}{\sigma_r}.$$

Note that in the case of a square, nonsingular matrix, this reduces to the definition (3.3).

The following perturbation theorem was proved by Wedin [106].

Theorem 6.12. *Assume that the matrix $A \in \mathbb{R}^{m\times n}$, where $m \geq n$ has full column rank, and let x be the solution of the least squares problem $\min_x \| Ax - b \|_2$. Let δA and δb be perturbations such that*

$$\eta = \frac{\| \delta A \|_2}{\sigma_n} = \kappa \epsilon_A < 1, \qquad \epsilon_A = \frac{\| \delta A \|_2}{\| A \|_2}.$$

Then the perturbed matrix $A + \delta A$ has full rank, and the perturbation of the solution δx satisfies

$$\| \delta x \|_2 \leq \frac{\kappa}{1-\eta} \left(\epsilon_A \| x \|_2 + \frac{\| \delta b \|_2}{\| A \|_2} + \epsilon_A \kappa \frac{\| r \|_2}{\| A \|_2} \right),$$

where r is the residual $r = b - Ax$.

There are at least two important observations to make here:

1. The number κ determines the condition of the least squares problem, and if $m = n$, then the residual r is equal to zero and the inequality becomes a perturbation result for a linear system of equations; cf. Theorem 3.5.

2. In the overdetermined case the residual is usually not equal to zero. Then the conditioning depends on κ^2. This dependence may be significant if the norm of the residual is large.

6.7 Rank-Deficient and Underdetermined Systems

Assume that A is rank-deficient, i.e., $\mathrm{rank}(A) = r < \min(m,n)$. The least squares problem can still be solved, but the solution is no longer unique. In this case we write the SVD

$$A = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix}, \tag{6.6}$$

where

$$U_1 \in \mathbb{R}^{m\times r}, \qquad \Sigma_1 \in \mathbb{R}^{r\times r}, \qquad V_1 \in \mathbb{R}^{n\times r}, \tag{6.7}$$

and the diagonal elements of Σ_1 are all nonzero. The norm of the residual can now be written

$$\| r \|_2^2 = \| Ax - b \|_2^2 = \left\| \begin{pmatrix} U_1 & U_2 \end{pmatrix} \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix} x - b \right\|_2^2.$$

Putting

$$y = V^T x = \begin{pmatrix} V_1^T x \\ V_2^T x \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \qquad \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} U_1^T b \\ U_2^T b \end{pmatrix}$$

and using the invariance of the norm under orthogonal transformations, the residual becomes

$$\| r \|_2^2 = \left\| \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right\|_2^2 = \| \Sigma_1 y_1 - b_1 \|_2^2 + \| b_2 \|_2^2.$$

Thus, we can minimize the residual by choosing $y_1 = \Sigma_1^{-1} b_1$. In fact,

$$y = \begin{pmatrix} \Sigma_1^{-1} b_1 \\ y_2 \end{pmatrix},$$

where y_2 is arbitrary, solves the least squares problem. Therefore, *the solution of the least squares problem is not unique*, and, since the columns of V_2 span the null-space of A, it is in this null-space, where the indeterminacy is. We can write

$$\| x \|_2^2 = \| y \|_2^2 = \| y_1 \|_2^2 + \| y_2 \|_2^2,$$

and therefore we obtain the solution of *minimum norm* by choosing $y_2 = 0$.

We summarize the derivation in a theorem.

Theorem 6.13 (minimum norm solution). *Assume that the matrix A is rank deficient with SVD* (6.6), (6.7). *Then the least squares problem* $\min_x \| Ax - b \|_2$ *does not have a unique solution. However, the problem*

$$\min_{x\in\mathcal{L}} \| x \|_2, \qquad \mathcal{L} = \{ x \mid \| Ax - b \|_2 = \min \},$$

has the unique solution

$$x = V \begin{pmatrix} \Sigma_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^T b = V_1 \Sigma_1^{-1} U_1^T b.$$

The matrix

$$A^\dagger = V \begin{pmatrix} \Sigma_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^T$$

is called the *pseudoinverse* of A. It is defined for any nonzero matrix of arbitrary dimensions.

The SVD can also be used to solve *underdetermined linear systems*, i.e., systems with more unknowns than equations. Let $A \in \mathbb{R}^{m\times n}$, with $m < n$, be given. The SVD of A is

$$A = U \begin{pmatrix} \Sigma & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix}, \qquad V_1 \in \mathbb{R}^{m\times m}. \tag{6.8}$$

Obviously A has full row rank if and only Σ is nonsingular.

We state a theorem concerning the solution of a linear system

$$Ax = b \tag{6.9}$$

for the case when A has full row rank.

Theorem 6.14 (solution of an underdetermined linear system). *Let $A \in \mathbb{R}^{m\times n}$ have full row rank with SVD* (6.8). *Then the linear system* (6.9) *always has a solution, which, however, is nonunique. The problem*

$$\min_{x\in\mathcal{K}} \| x \|_2, \qquad \mathcal{K} = \{x \mid Ax = b\}, \tag{6.10}$$

has the unique solution

$$x = V_1 \Sigma^{-1} U^T b. \tag{6.11}$$

Proof. Using the SVD (6.8) we can write

$$Ax = U \begin{pmatrix} \Sigma & 0 \end{pmatrix} \begin{pmatrix} V_1^T x \\ V_2^T x \end{pmatrix} =: U \begin{pmatrix} \Sigma & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = U\Sigma y_1.$$

Since Σ is nonsingular, we see that for any right-hand side, (6.11) is a solution of the linear system. However, we can add an arbitrary solution component in the null-space of A, $y_2 = V_2^T x$, and we still have a solution. The minimum norm solution, i.e., the solution of (6.10), is given by (6.11). □

The rank-deficient case may or may not have a solution depending on the right-hand side, and that case can be easily treated as in Theorem 6.13.

6.8 Computing the SVD

The SVD is computed in MATLAB by the statement `[U,S,V]=svd(A)`. This statement is an implementation of algorithms from LAPACK [1]. (The double precision high-level driver algorithm for SVD is called `DGESVD`.) In the algorithm the matrix is first reduced to bidiagonal form by a series of Householder transformations from the left and right. Then the bidiagonal matrix is iteratively reduced to diagonal form using a variant of the QR algorithm; see Chapter 15.

The SVD of a dense (full) matrix can be computed in $\mathcal{O}(mn^2)$ flops. Depending on how much is computed, the constant is of the order 5–25.

The computation of a partial SVD of a large, sparse matrix is done in MATLAB by the statement `[U,S,V]=svds(A,k)`. This statement is based on Lanczos methods from ARPACK. We give a brief description of Lanczos algorithms in Chapter 15. For a more comprehensive treatment, see [4].

6.9 Complete Orthogonal Decomposition

In the case when the matrix is rank deficient, computing the SVD is the most reliable method for determining the rank. However, it has the drawbacks that it is comparatively expensive to compute, and it is expensive to update (when new rows and/or columns are added). Both these issues may be critical, e.g., in a real-time application. Therefore, methods have been developed that approximate the SVD, so-called *complete orthogonal decompositions*, which in the noise-free case and in exact arithmetic can be written

$$A = Q \begin{pmatrix} T & 0 \\ 0 & 0 \end{pmatrix} Z^T$$

for orthogonal Q and Z and triangular $T \in \mathbb{R}^{r \times r}$ when A has rank r. Obviously, the SVD is a special case of a complete orthogonal decomposition.

In this section we will assume that the matrix $A \in \mathbb{R}^{m \times n}$, $m \geq n$, has exact or numerical rank r. (Recall the definition of numerical rank on p. 63.)

6.9.1 QR with Column Pivoting

The first step toward obtaining a complete orthogonal decomposition is to perform column pivoting in the computation of a QR decomposition [22]. Consider the matrix before the algorithm has started: compute the 2-norm of each column, and move the column with largest norm to the leftmost position. This is equivalent to multiplying the matrix by a permutation matrix P_1 from the right. In the first step of the reduction to triangular form, a Householder transformation is applied that annihilates elements in the first column:

$$A \longrightarrow AP_1 \longrightarrow Q_1^T AP_1 = \begin{pmatrix} r_{11} & r_1^T \\ 0 & B \end{pmatrix}.$$

Then in the next step, find the column with largest norm in B, permute the columns so that the column with largest norm is moved to position 2 in A (this involves only columns 2 to n, of course), and reduce the first column of B:

$$Q_1^T A P_1 \longrightarrow Q_2^T Q_1^T A P_1 P_2 = \begin{pmatrix} r_{11} & r_{12} & \bar{r}_1^T \\ 0 & r_{22} & \bar{r}_2^T \\ 0 & 0 & C \end{pmatrix}.$$

It is seen that $|r_{11}| \geq |r_{22}|$.

After n steps we have computed

$$AP = Q \begin{pmatrix} R \\ 0 \end{pmatrix}, \qquad Q = Q_1 Q_2 \cdots Q_n, \qquad P = P_1 P_2 \cdots P_{n-1}.$$

The product of permutation matrices is itself a permutation matrix.

Proposition 6.15. *Assume that A has rank r. Then, in exact arithmetic, the QR decomposition with column pivoting is given by*

$$AP = Q \begin{pmatrix} R_{11} & R_{12} \\ 0 & 0 \end{pmatrix}, \qquad R_{11} \in \mathbb{R}^{r \times r},$$

and the diagonal elements of R_{11} are nonzero (R_{11} is nonsingular).

Proof. Obviously the diagonal elements occurring in the process are nonincreasing:

$$|r_{11}| \geq |r_{22}| \geq \cdots .$$

Assume that $r_{r+1,r+1} > 0$. That would imply that the rank of R in the QR decomposition is larger than r, which is a contradiction, since R and A must have the same rank. □

Example 6.16. The following MATLAB script performs QR decomposition with column pivoting on a matrix that is constructed to be rank deficient:

```
[U,ru]=qr(randn(3)); [V,rv]=qr(randn(3));
D=diag([1 0.5 0]);  A=U*D*V';
[Q,R,P]=qr(A);   % QR with column pivoting

R = -0.8540    0.4311   -0.0642
          0    0.4961   -0.2910
          0         0    0.0000

>> R(3,3) = 2.7756e-17
```

In many cases QR decomposition with pivoting gives reasonably accurate information about the numerical rank of a matrix. We modify the matrix in the previous script by adding noise:

```
[U,ru]=qr(randn(3)); [V,rv]=qr(randn(3));
D=diag([1 0.5 0]);  A=U*D*V'+1e-4*randn(3);
[Q,R,P]=qr(A);

>> R =  0.8172   -0.4698   -0.1018
             0   -0.5758    0.1400
             0         0    0.0001
```

The smallest diagonal element is of the same order of magnitude as the smallest singular value:

```
>> svd(A) = 1.0000
            0.4999
            0.0001
```

■

It turns out, however, that one cannot rely completely on this procedure to give correct information about possible rank deficiency. We give an example due to Kahan; see [50, Section 8.3].

Example 6.17. Let $c^2 + s^2 = 1$. For n large enough, the triangular matrix

$$T_n(c) = \mathrm{diag}(1, s, s^2, \dots, s^{n-1}) \begin{pmatrix} 1 & -c & -c & \cdots & -c \\ & 1 & -c & \cdots & -c \\ & & \ddots & & \vdots \\ & & & 1 & -c \\ & & & & 1 \end{pmatrix}$$

is very ill-conditioned. For $n = 200$ and $c = 0.2$, we have

$$\kappa_2(T_n(c)) = \frac{\sigma_1}{\sigma_n} \approx \frac{12.7}{5.7 \cdot 10^{-18}}.$$

Thus, in IEEE double precision, the matrix is singular. The columns of the triangular matrix all have length 1. Therefore, because the elements in each row to the right of the diagonal are equal, QR decomposition with column pivoting will not introduce any column interchanges, and the upper triangular matrix R is equal to $T_n(c)$. However, the bottom diagonal element is equal to $s^{199} \approx 0.0172$, so for this matrix QR with column pivoting does not give any information whatsoever about the ill-conditioning. ■