Positive Definite Matrices

6.3 Singular Value Decomposition

6.3 Singular Value Decomposition

A great matrix factorization has been saved for the end of the basic course. $U\Sigma V^{\mathrm{T}}$ joins with LU from elimination and QR from orthogonalization (Gauss and Gram-Schmidt). Nobody's name is attached; $A = U\Sigma V^{\mathrm{T}}$ is known as the "SVD" or the ***singular value decomposition***. We want to describe it, to prove it, and to discuss its applications—which are many and growing.

The SVD is closely associated with the eigenvalue-eigenvector factorization $Q\Lambda Q^{\mathrm{T}}$ of a positive definite matrix. The eigenvalues are in the diagonal matrix Λ. The eigenvector matrix Q *is orthogonal* ($Q^{\mathrm{T}}Q = I$) because eigenvectors of a symmetric matrix can be chosen to be orthonormal. For most matrices that is not true, and for rectangular matrices it is ridiculous (eigenvalues undefined). But now we allow the Q on the left and the Q^{T} on the right to be *any two orthogonal matrices* U and V^{T}—not necessarily transposes of each other. Then every matrix will split into $A = U\Sigma V^{\mathrm{T}}$.

The diagonal (but rectangular) matrix Σ has eigenvalues from $A^{\mathrm{T}}A$, not from A! Those positive entries (also called sigma) will be $\sigma_1, \ldots, \sigma_r$. They are the ***singular values*** of A. They fill the first r places on the main diagonal of Σ—when A has rank r. The rest of Σ is zero.

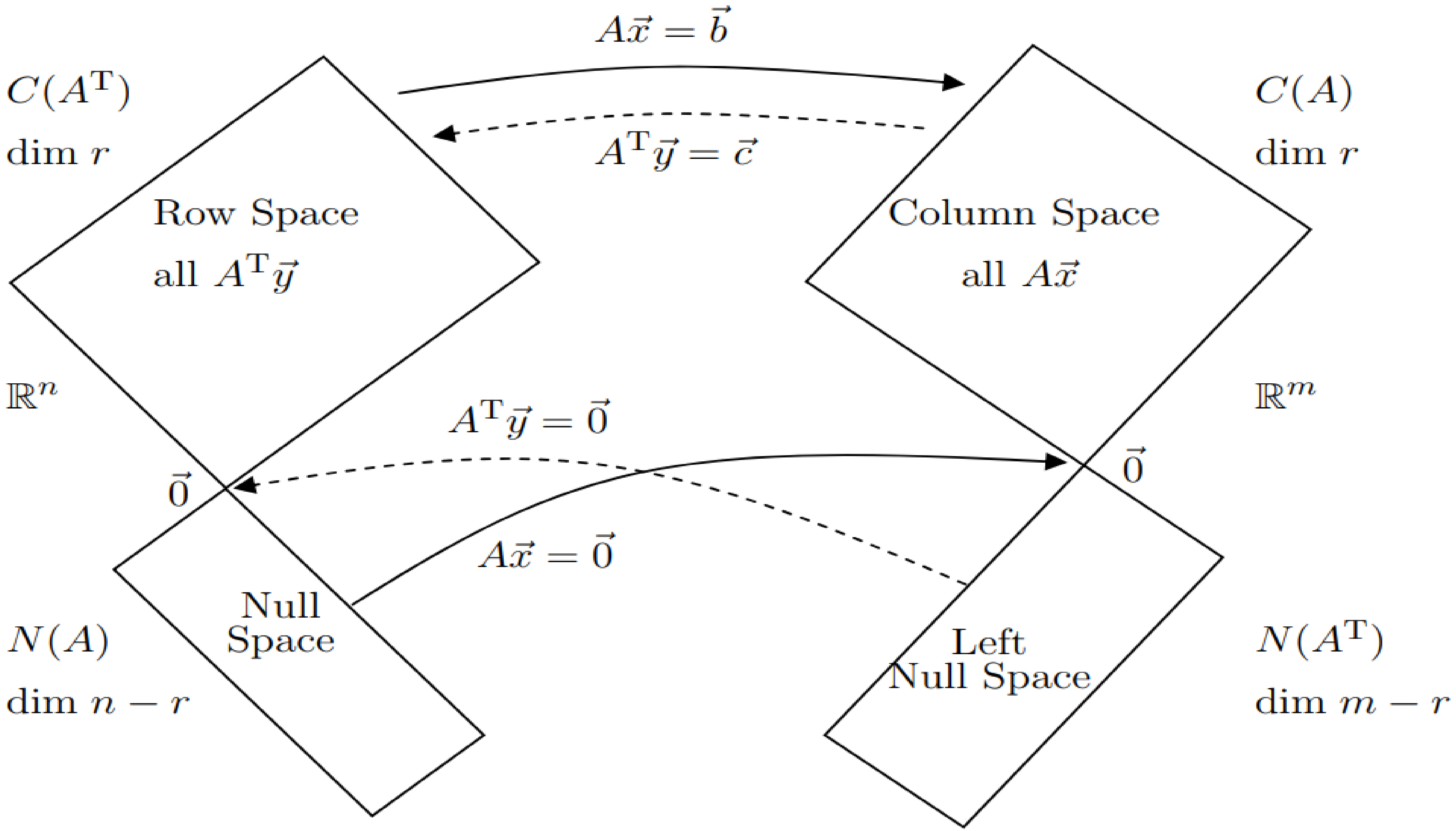
With rectangular matrices, the key is almost always to consider $A^{\mathrm{T}}A$ and AA^{T}.

Singular Value Decomposition: Any m by n matrix A can be factored into

$$A = U\Sigma V^{\mathrm{T}} = (\textbf{orthogonal})(\textbf{diagonal})(\textbf{orthogonal}).$$

The columns of U (m by m) are eigenvectors of AA^{T}, and the columns of V (n by n) are eigenvectors of $A^{\mathrm{T}}A$. The r singular values on the diagonal of Σ (m by n) are the square roots of the nonzero eigenvalues of both AA^{T} and $A^{\mathrm{T}}A$.

***Remark* 1.** For positive definite matrices, Σ is Λ and $U\Sigma V^{\mathrm{T}}$ is identical to $Q\Lambda Q^{\mathrm{T}}$. For other symmetric matrices, any negative eigenvalues in Λ become positive in Σ. For complex matrices, Σ remains real but U and V become *unitary* (the complex version of orthogonal). We take complex conjugates in $U^{\mathrm{H}}U = I$ and $V^{\mathrm{H}}V = I$ and $A = U\Sigma V^{\mathrm{H}}$.


$A\vec{x} = \vec{b}$
$C(A^{\mathrm{T}})$
dim r
$A^{\mathrm{T}}\vec{y} = \vec{c}$
$C(A)$
dim r
Row Space
all $A^{\mathrm{T}}\vec{y}$
Column Space
all $A\vec{x}$
$\mathbb{R}^n$
$\mathbb{R}^m$
$A^{\mathrm{T}}\vec{y} = \vec{0}$
$\vec{0}$
$\vec{0}$
$A\vec{x} = \vec{0}$
Null Space
$N(A)$
dim $n - r$
Left Null Space
$N(A^{\mathrm{T}})$
dim $m - r$

***Remark* 2.** U and V give orthonormal bases for *all four fundamental subspaces*:

first	r	columns of U:	**column space** of A
last	$m-r$	columns of U:	**left nullspace** of A
first	r	columns of V:	**row space** of A
last	$n-r$	columns of V:	**nullspace** of A

***Remark* 3.** The SVD chooses those bases in an extremely special way. They are more than just orthonormal. *When A multiplies a column v_j of V, it produces σ_j times a column of U*. That comes directly from $AV = U\Sigma$, looked at a column at a time.

***Remark* 4.** Eigenvectors of AA^{T} and $A^{\mathrm{T}}A$ must go into the columns of U and V:

$$AA^{\mathrm{T}} = (U\Sigma V^{\mathrm{T}})(V\Sigma^{\mathrm{T}}U^{\mathrm{T}}) = U\Sigma\Sigma^{\mathrm{T}}U^{\mathrm{T}} \quad \text{and, similarly,} \quad A^{\mathrm{T}}A = V\Sigma^{\mathrm{T}}\Sigma V^{\mathrm{T}}. \tag{1}$$

U must be the eigenvector matrix for AA^{T}. The eigenvalue matrix in the middle is $\Sigma\Sigma^{\mathrm{T}}$—which is m by m with $\sigma_1^2, \ldots, \sigma_r^2$ on the diagonal.

From the $A^{\mathrm{T}}A = V\Sigma^{\mathrm{T}}\Sigma V^{\mathrm{T}}$, the *$V$ matrix must be the eigenvector matrix for $A^{\mathrm{T}}A$*. The diagonal matrix $\Sigma^{\mathrm{T}}\Sigma$ has the same $\sigma_1^2, \ldots, \sigma_r^2$, but it is n by n.

***Remark* 5.** Here is the reason that $Av_j = \sigma_j u_j$. Start with $A^{\mathrm{T}}Av_j = \sigma_j^2 v_j$:

$$\textbf{Multiply by } A \qquad AA^{\mathrm{T}}Av_j = \sigma_j^2 Av_j \tag{2}$$

This says that Av_j is an eigenvector of AA^{T}! We just moved parentheses to $(AA^{\mathrm{T}})(Av_j)$. The length of this eigenvector Av_j is σ_j, because

$$v^{\mathrm{T}}A^{\mathrm{T}}Av_j = \sigma_j^2 v_j^{\mathrm{T}} v_j \qquad \text{gives} \qquad \|Av_j\|^2 = \sigma_j^2.$$

So the unit eigenvector is $Av_j/\sigma_j = u_j$. **In other words,** $AV = U\Sigma$.

Example 1. This A has only one column: rank $r = 1$. Then Σ has only $\sigma_1 = 3$:

$$\textbf{SVD} \qquad A = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix} = U_{3\times 3}\Sigma_{3\times 1}V_{1\times 1}^{\mathrm{T}}.$$

$A^{\mathrm{T}}A$ is 1 by 1, whereas AA^{T} is 3 by 3. They both have eigenvalue 9 (whose square root is the 3 in Σ). The two zero eigenvalues of AA^{T} leave some freedom for the eigenvectors in columns 2 and 3 of U. We kept that matrix orthogonal.

Example 2. Now A has rank 2, and $AA^{\mathrm{T}} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ with $\lambda = 3$ and 1:

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = U\Sigma V^{\mathrm{T}} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{matrix} /\sqrt{6} \\ /\sqrt{2} \\ /\sqrt{3} \end{matrix}.$$

Notice $\sqrt{3}$ and $\sqrt{1}$. The columns of U are *left* singular vectors (unit eigenvectors of AA^{T}). The columns of V are *right* singular vectors (unit eigenvectors of $A^{\mathrm{T}}A$).

Problem Set 6.3

Problems 1–2 compute the SVD of a square singular matrix A.

1. Compute $A^{\mathrm{T}}A$ and its eigenvalues σ_1^2, 0 and unit eigenvectors v_1, v_2:

$$A = \begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix}.$$

2. (a) Compute AA^{T} and its eigenvalues σ_1^2, 0 and unit eigenvectors u_1, u_2.

(b) Choose signs so that $Av_1 = \sigma_1 u_1$ and verify the SVD:

$$\begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix} = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & 0 \end{bmatrix} \begin{bmatrix} v_1 & v_2 \end{bmatrix}^{\mathrm{T}}.$$

(c) Which four vectors give orthonormal bases for $\boldsymbol{C}(A)$, $\boldsymbol{N}(A)$, $\boldsymbol{C}(A^{\mathrm{T}})$, $\boldsymbol{N}(A^{\mathrm{T}})$?

5. Compute $A^{\mathrm{T}}A$ and AA^{T}, and their eigenvalues and unit eigenvectors, for

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Multiply the three matrices $U\Sigma V^{\mathrm{T}}$ to recover A.

6. Suppose $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ are orthonormal bases for $\mathbf{R}^n$. Construct the matrix A that transforms each v_j into u_j to give $Av_1 = u_1, \ldots, Av_n = u_n$.

7. Construct the matrix with rank 1 that has $Av = 12u$ for $v = \frac{1}{2}(1,1,1,1)$ and $u = \frac{1}{3}(2,2,1)$. Its only singular value is $\sigma_1 =$ ____.

8. Find $U\Sigma V^{\mathrm{T}}$ if A has orthogonal columns $w_1, \ldots, w_n$ of lengths $\sigma_1, \ldots, \sigma_n$.

10. Suppose A is a 2 by 2 symmetric matrix with unit eigenvectors u_1 and u_2. If its eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = -2$, what are U, Σ, and V^{T}?

11. Suppose A is invertible (with $\sigma_1 > \sigma_2 > 0$). Change A by as small a matrix as possible to produce a *singular* matrix A_0. *Hint*: U and V do not change:

$$\text{Find } A_0 \text{ from} \qquad A = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & \sigma_2 \end{bmatrix} \begin{bmatrix} v_1 & v_2 \end{bmatrix}^{\mathrm{T}}.$$

12. (a) If A changes to $4A$, what is the change in the SVD?

(b) What is the SVD for A^{T} and for A^{-1}?