LECTURE 20

The singular value decomposition

Every matrix is a rotation, a stretch, and another rotation

Dr Ilia Pirashvili
Department of Mathematics, University of Galway
11 Novemeber 2026 · Winter 2026

Where we got to

LAST LECTURE

- $A\mathbf{v} = \lambda\mathbf{v}$: the eigenvalues are the roots of $\mathrm{Det}(A - \lambda I)$, the eigenspaces are $\mathrm{Ker}(A - \lambda I)$.
- $A = EDE^{-1}$ exactly when A has n independent eigenvectors.
- Then $A^k = ED^kE^{-1}$, and the eigenvalues decide the long run.

Diagonalisation needs a square matrix, and enough eigenvectors. Today we build a factorisation that works for *every* matrix, square or not, invertible or not, and exposes exactly what it does to space.

Warm-up

WARM-UP

Compute $A^{\mathsf{T}}A$ for $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$. What do you notice about the result?

Warm-up

WARM-UP

Compute $A^{\mathsf{T}}A$ for $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$. What do you notice about the result?

SOLUTION

$A^{\mathsf{T}}A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, which is square ($2 \times 2$, not 3×3) and *symmetric*. Both are automatic, since $(A^{\mathsf{T}}A)^{\mathsf{T}} = A^{\mathsf{T}}(A^{\mathsf{T}})^{\mathsf{T}} = A^{\mathsf{T}}A$.

That symmetry is the seed of this lecture.

SECTION 1

Symmetric matrices

Symmetric matrices behave beautifully

THEOREM 20.1

Let $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ be symmetric, that is $A^{\mathsf{T}} = A$. Then

1. *every eigenvalue of A is real;*
2. *eigenvectors for distinct eigenvalues are orthogonal;*
3. *$\mathbb{R}^n$ has an orthonormal basis of eigenvectors of A.*

We take these on trust; (3) is the *spectral theorem*.

Symmetric matrices behave beautifully

THEOREM 20.1

Let $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ be symmetric, that is $A^{\mathsf{T}} = A$. Then

1. *every eigenvalue of A is real;*
2. *eigenvectors for distinct eigenvalues are orthogonal;*
3. *$\mathbb{R}^n$ has an orthonormal basis of eigenvectors of A.*

We take these on trust; (3) is the *spectral theorem*.

EXAMPLE 20.2

$A = \left[\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right]$ had eigenvectors $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$, orthogonal as promised. Normalising gives the orthonormal basis $\frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right], \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$.

Arguing from the eigenvector equation

To find out what an eigenvalue can be, start from $A\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, and apply what you know about A.

Arguing from the eigenvector equation

To find out what an eigenvalue can be, start from $A\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, and apply what you know about A.

EXAMPLE 20.3

If $P^2 = P$ then every eigenvalue of P is 0 or 1. Apply P again:

$$\lambda\mathbf{v} = P\mathbf{v} = P^2\mathbf{v} = P(\lambda\mathbf{v}) = \lambda P\mathbf{v} = \lambda^2\mathbf{v},$$

so $(\lambda^2 - \lambda)\mathbf{v} = \mathbf{0}$. As $\mathbf{v} \neq \mathbf{0}$, we get $\lambda^2 = \lambda$.

Arguing from the eigenvector equation

To find out what an eigenvalue can be, start from $A\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, and apply what you know about A.

EXAMPLE 20.3

If $P^2 = P$ then every eigenvalue of P is 0 or 1. Apply P again:

$$\lambda\mathbf{v} = P\mathbf{v} = P^2\mathbf{v} = P(\lambda\mathbf{v}) = \lambda P\mathbf{v} = \lambda^2\mathbf{v},$$

so $(\lambda^2 - \lambda)\mathbf{v} = \mathbf{0}$. As $\mathbf{v} \neq \mathbf{0}$, we get $\lambda^2 = \lambda$.

EXAMPLE 20.4

If $\|A\mathbf{x}\| \leq c\|\mathbf{x}\|$ for all $\mathbf{x}$, then every real eigenvalue has $|\lambda| \leq c$. Take lengths:

$$|\lambda|\,\|\mathbf{v}\| = \|\lambda\mathbf{v}\| = \|A\mathbf{v}\| \leq c\|\mathbf{v}\|,$$

and divide by $\|\mathbf{v}\| > 0$.

The eigenvalues of $A^{\mathsf{T}}A$ are never negative

THEOREM 20.5

Let $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$. Every eigenvalue of $A^{\mathsf{T}}A$ is real and ≥ 0.

The eigenvalues of $A^{\mathsf{T}}A$ are never negative

THEOREM 20.5

Let $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$. Every eigenvalue of $A^{\mathsf{T}}A$ is real and ≥ 0.

PROOF.

$A^{\mathsf{T}}A$ is symmetric, so its eigenvalues are real. Let $A^{\mathsf{T}}A\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$, and multiply on the left by $\mathbf{v}^{\mathsf{T}}$:

$$\mathbf{v}^{\mathsf{T}}A^{\mathsf{T}}A\,\mathbf{v} \;=\; \lambda\,\mathbf{v}^{\mathsf{T}}\mathbf{v} \qquad \Longrightarrow \qquad (A\mathbf{v})^{\mathsf{T}}(A\mathbf{v}) \;=\; \lambda\|\mathbf{v}\|^2,$$

that is $\|A\mathbf{v}\|^2 = \lambda\|\mathbf{v}\|^2$. Hence

$$\lambda \;=\; \frac{\|A\mathbf{v}\|^2}{\|\mathbf{v}\|^2} \;\geq\; 0,$$

a ratio of two non-negative numbers with positive denominator. □

SECTION 2

Singular values

Definition

DEFINITION 20.6 · Singular values

Let $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ and let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ be the eigenvalues of $A^{\mathsf{T}}A$, listed with multiplicity in decreasing order. The **singular values** of A are

$$s_i \coloneqq \sqrt{\lambda_i}, \qquad \text{so} \qquad s_1 \geq s_2 \geq \cdots \geq s_n \geq 0.$$

The square roots are real because the λ_i are non-negative, which is precisely the theorem we just proved. Ordering them is a convention, but a universal one.

Definition

DEFINITION 20.6 · Singular values

Let $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ and let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ be the eigenvalues of $A^{\mathsf{T}}A$, listed with multiplicity in decreasing order. The **singular values** of A are

$$s_i \;:=\; \sqrt{\lambda_i}, \qquad \text{so} \qquad s_1 \geq s_2 \geq \cdots \geq s_n \geq 0.$$

The square roots are real because the λ_i are non-negative, which is precisely the theorem we just proved. Ordering them is a convention, but a universal one.

From the proof, $s_i = \|A\mathbf{v}_i\|$ for a *unit* eigenvector $\mathbf{v}_i$: a singular value measures how far A stretches that direction.

A first computation

EXAMPLE 20.7

Find the singular values of $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

A first computation

EXAMPLE 20.7

Find the singular values of $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

$$A^{\mathsf{T}}A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}.$$

A first computation

EXAMPLE 20.7

Find the singular values of $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

$$A^{\mathsf{T}}A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}.$$

$$\mathrm{Det}(A^{\mathsf{T}}A - \lambda I) = (2-\lambda)^2 - 4 = \lambda^2 - 4\lambda = \lambda(\lambda - 4),$$

so $\lambda_1 = 4$ and $\lambda_2 = 0$.

A first computation

EXAMPLE 20.7

Find the singular values of $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

$$A^{\mathsf{T}}A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}.$$

$$\mathrm{Det}(A^{\mathsf{T}}A - \lambda I) = (2-\lambda)^2 - 4 = \lambda^2 - 4\lambda = \lambda(\lambda - 4),$$

so $\lambda_1 = 4$ and $\lambda_2 = 0$.

$$s_1 = 2, \qquad s_2 = 0.$$

A zero singular value signals a collapsed direction: here $\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$ is killed by A, and Rank $A = 1$.

A rectangular one

EXAMPLE 20.8

Find the singular values of $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$.

A rectangular one

EXAMPLE 20.8

Find the singular values of $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$.

From the warm-up, $A^{\mathsf{T}}A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, whose eigenvalues we computed last lecture: 3 and 1.

So $s_1 = \sqrt{3}$ and $s_2 = 1$.

A rectangular one

EXAMPLE 20.8

Find the singular values of $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$.

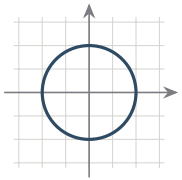
From the warm-up, $A^{\mathsf{T}}A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, whose eigenvalues we computed last lecture: 3 and 1.

So $s_1 = \sqrt{3}$ and $s_2 = 1$.

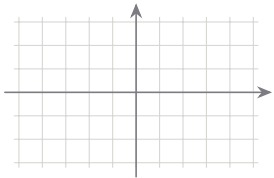
A is 3×2, yet there are only two singular values: always n of them, one per *column*, since $A^{\mathsf{T}}A$ is $n \times n$. The number of *non-zero* ones will turn out to be Rank A.

What the singular values are telling us

Apply A to the unit circle of $\mathbb{R}^2$. The image is an ellipse, and its semi-axes have lengths s_1 and s_2.



the unit circle



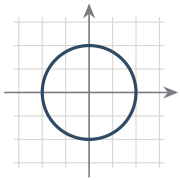
its image under A

NOT EXAMINABLE

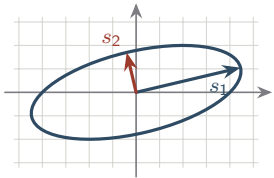
What the singular values are telling us

Apply A to the unit circle of $\mathbb{R}^2$. The image is an ellipse, and its semi-axes have lengths s_1 and s_2.

For $A = \left[\begin{smallmatrix} 2 & 1 \\ 0 & 1 \end{smallmatrix}\right]$, the singular values are about 2.29 and 0.87: the circle is stretched a little over twice in one direction and squashed slightly in another.



the unit circle



its image under A

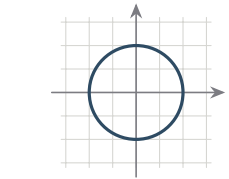
NOT EXAMINABLE

What the singular values are telling us

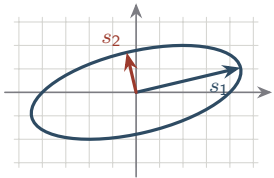
Apply A to the unit circle of $\mathbb{R}^2$. The image is an ellipse, and its semi-axes have lengths s_1 and s_2.

For $A = \left[\begin{smallmatrix} 2 & 1 \\ 0 & 1 \end{smallmatrix}\right]$, the singular values are about 2.29 and 0.87: the circle is stretched a little over twice in one direction and squashed slightly in another.

So s_1 is the greatest stretch A can apply to any unit vector, and s_n the least. That is the whole geometric content of the definition.



the unit circle



its image under A

NOT EXAMINABLE

SECTION 3

The theorem

Statement

THEOREM 20.9 · Singular value decomposition

Let $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ *with singular values* $s_1 \geq \cdots \geq s_n \geq 0$. *Then there are orthogonal matrices* $U \in \mathrm{Mat}_{m\times m}(\mathbb{R})$ *and* $V \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ *with*

$$A \;=\; U\Sigma V^{\mathsf{T}},$$

where $\Sigma \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ *has the* s_i *down its leading diagonal (as many as fit) and zeros elsewhere.*

Statement

THEOREM 20.9 · Singular value decomposition

Let $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ *with singular values* $s_1 \geq \cdots \geq s_n \geq 0$. *Then there are orthogonal matrices* $U \in \mathrm{Mat}_{m\times m}(\mathbb{R})$ *and* $V \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ *with*

$$A \;=\; U\Sigma V^{\mathsf{T}},$$

where $\Sigma \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ *has the* s_i *down its leading diagonal (as many as fit) and zeros elsewhere.*

The columns satisfy

$$A^{\mathsf{T}}A\,\mathbf{v}_i \;=\; s_i^2\,\mathbf{v}_i \qquad \text{and} \qquad A\,\mathbf{v}_i \;=\; s_i\,\mathbf{u}_i,$$

where $\mathbf{v}_i = v_{-,i}$ and $\mathbf{u}_i = u_{-,i}$.

The shape of Σ

Σ has the same shape as A, so it is usually not square:

$$m > n: \quad \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad m < n: \quad \begin{bmatrix} s_1 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 \end{bmatrix}$$

The shape of Σ

Σ has the same shape as A, so it is usually not square:

$$m > n: \quad \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad m < n: \quad \begin{bmatrix} s_1 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 \end{bmatrix}$$

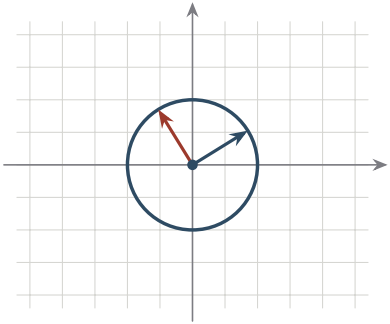
KEY IDEA

Read $A = U\Sigma V^{\mathsf{T}}$ from the right: V^{T} rotates the source, Σ stretches each axis by s_i, and U rotates the target. *Every matrix is a rotation, a stretch along axes, and another rotation*, and nothing else ever happens.

Here a "rotation" is any orthogonal matrix, so it may include a reflection.

What that looks like

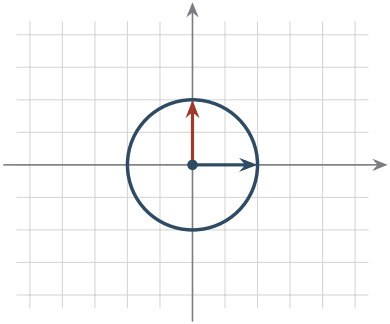
Start with the unit circle in the source, and two perpendicular radii to keep track of.



x

What that looks like

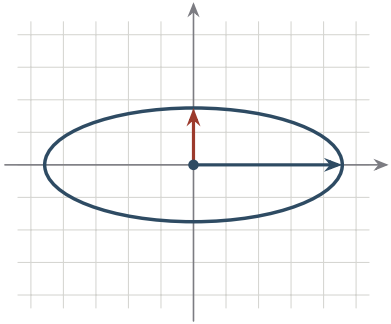
V^{T} rotates. A circle is still a circle, but the two marked radii now lie along the axes.



$V^{\mathsf{T}}\mathbf{x}$

What that looks like

Σ stretches the axes by s_1 and s_2. The circle becomes an ellipse with those semi-axes.

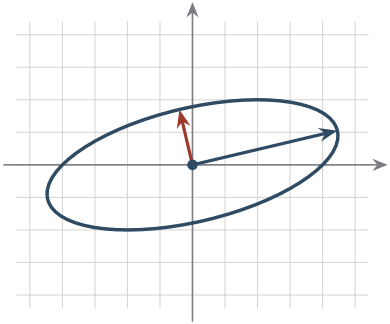


$\Sigma V^{\mathsf{T}}\mathbf{x}$

What that looks like

U rotates once more. The ellipse turns into its final position.

The whole action of A was: *turn*, *stretch along axes*, *turn*.



$$U\Sigma V^{\mathsf{T}}\mathbf{x} = A\mathbf{x}$$

Why it works

A sketch, in three steps.

- $A^{\mathsf{T}}A$ is symmetric, so by the spectral theorem $\mathbb{R}^n$ has an *orthonormal* basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of eigenvectors, with eigenvalues $s_i^2 \geq 0$. Take V to have these as columns; it is orthogonal because its columns are orthonormal.

NOT EXAMINABLE

Why it works

A sketch, in three steps.

- $A^\mathsf{T}A$ is symmetric, so by the spectral theorem $\mathbb{R}^n$ has an *orthonormal* basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of eigenvectors, with eigenvalues $s_i^2 \geq 0$. Take V to have these as columns; it is orthogonal because its columns are orthonormal.
- For each i with $s_i > 0$, set $\mathbf{u}_i = A\mathbf{v}_i / s_i$. These are orthonormal, since

$$\mathbf{u}_i \cdot \mathbf{u}_j = \frac{\mathbf{v}_i^\mathsf{T} A^\mathsf{T} A \mathbf{v}_j}{s_i s_j} = \frac{s_j^2}{s_i s_j} \mathbf{v}_i \cdot \mathbf{v}_j,$$

  which is 1 if $i = j$ and 0 otherwise.

NOT EXAMINABLE

Why it works

A sketch, in three steps.

- $A^{\mathsf{T}}A$ is symmetric, so by the spectral theorem $\mathbb{R}^n$ has an *orthonormal* basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of eigenvectors, with eigenvalues $s_i^2 \geq 0$. Take V to have these as columns; it is orthogonal because its columns are orthonormal.
- For each i with $s_i > 0$, set $\mathbf{u}_i = A\mathbf{v}_i / s_i$. These are orthonormal, since

$$\mathbf{u}_i \cdot \mathbf{u}_j = \frac{\mathbf{v}_i^{\mathsf{T}} A^{\mathsf{T}} A \mathbf{v}_j}{s_i s_j} = \frac{s_j^2}{s_i s_j} \mathbf{v}_i \cdot \mathbf{v}_j,$$

  which is 1 if $i = j$ and 0 otherwise.
- Extend $\mathbf{u}_1, \ldots, \mathbf{u}_r$ to an orthonormal basis of $\mathbb{R}^m$, using Gram–Schmidt, and let U have those columns. Then $AV = U\Sigma$, which is $A = U\Sigma V^{\mathsf{T}}$.

NOT EXAMINABLE

SECTION 4

Computing an SVD

The recipe

1. Form $A^{\mathsf{T}}A$, an $n \times n$ symmetric matrix.
2. Find its eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$, and set $s_i = \sqrt{\lambda_i}$.
3. For each λ_i, find an orthonormal basis of $\mathrm{Ker}(A^{\mathsf{T}}A - \lambda_i I)$. Collect these as the columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of V, in the same order.
4. Build Σ: the same shape as A, with the s_i on the diagonal.
5. For each i with $s_i > 0$, set $\mathbf{u}_i = \dfrac{1}{s_i} A\mathbf{v}_i$.
6. If fewer than m columns have appeared, extend $\mathbf{u}_1, \ldots, \mathbf{u}_r$ to an orthonormal basis of $\mathbb{R}^m$.

Steps 3 and 6 are where Gram–Schmidt earns its keep.

Notes on the steps

ORDER AND LENGTH

The $\mathbf{v}_i$ must be *unit* vectors, listed in the *same* order as the s_i. Otherwise the factorisation breaks.

Notes on the steps

ORDER AND LENGTH

The $\mathbf{v}_i$ must be *unit* vectors, listed in the *same* order as the s_i. Otherwise the factorisation breaks.

REPEATED EIGENVALUES

If some λ is repeated, its eigenspace has dimension > 1 and the eigenvectors it hands you need not be orthogonal. Apply Gram–Schmidt *within* that eigenspace.

Notes on the steps

ORDER AND LENGTH

The $\mathbf{v}_i$ must be *unit* vectors, listed in the *same* order as the s_i. Otherwise the factorisation breaks.

REPEATED EIGENVALUES

If some λ is repeated, its eigenspace has dimension > 1 and the eigenvectors it hands you need not be orthogonal. Apply Gram–Schmidt *within* that eigenspace.

Two cheap checks: $V^{\mathsf{T}}V = I$ and $U^{\mathsf{T}}U = I$; then multiply out $U\Sigma V^{\mathsf{T}}$ and compare with A.

Worked example 1 (1/3)

Find an SVD of $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$, a 3×2 matrix.

Worked example 1 (1/3)

Find an SVD of $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$, a 3×2 matrix.

Steps 1–2. $A^{\mathsf{T}}A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ with eigenvalues 3 and 1, so

$$s_1 = \sqrt{3}, \qquad s_2 = 1.$$

Worked example 1 (1/3)

Find an SVD of $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$, a 3×2 matrix.

Steps 1–2. $A^{\mathsf{T}}A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ with eigenvalues 3 and 1, so

$$s_1 = \sqrt{3}, \qquad s_2 = 1.$$

Step 3. $\lambda = 3$: kernel of $\left[\begin{smallmatrix} -1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$, spanned by $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$. $\lambda = 1$: kernel of $\left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]$, spanned by $\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$. Normalising,

$$V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Worked example 1 (2/3)

Step 4. Σ is 3×2:

$$\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Worked example 1 (2/3)

Step 4. Σ is 3×2:

$$\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Step 5.

$$\mathbf{u}_1 = \frac{1}{\sqrt{3}} A\mathbf{v}_1 = \frac{1}{\sqrt{3}} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \qquad \mathbf{u}_2 = \frac{1}{1} A\mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

Worked example 1 (2/3)

Step 4. Σ is 3×2:

$$\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Step 5.

$$\mathbf{u}_1 = \frac{1}{\sqrt{3}} A\mathbf{v}_1 = \frac{1}{\sqrt{3}} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \qquad \mathbf{u}_2 = \frac{1}{1} A\mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

Check they are orthonormal: $\mathbf{u}_1 \cdot \mathbf{u}_2 = \frac{1}{\sqrt{12}}(1 - 1 + 0) = 0$, and each has length 1.

The reduced form

When $m > n$, the bottom rows of Σ are zero, so the last $m - n$ columns of U only ever get multiplied by zeros. We may as well leave them out:

$$A = \underbrace{\begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{6} & -1/\sqrt{2} \\ 2/\sqrt{6} & 0 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix}}_{\Sigma} \underbrace{\tfrac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{V^{\mathsf{T}}}.$$

The reduced form

When $m > n$, the bottom rows of Σ are zero, so the last $m - n$ columns of U only ever get multiplied by zeros. We may as well leave them out:

$$A = \underbrace{\begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{6} & -1/\sqrt{2} \\ 2/\sqrt{6} & 0 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix}}_{\Sigma} \underbrace{\tfrac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{V^{\mathsf{T}}}.$$

THE REDUCED SVD

For $m > n$ and $\operatorname{Rank} A = n$, the **reduced SVD** has U of size $m \times n$ and Σ of size $n \times n$. It needs steps 1 to 5 only.

Now $U^{\mathsf{T}}U = I_n$, but U is not square. Many books state the SVD in this form, so always look at the size of U you are asked for.

Worked example 1 (3/3): the full U

Step 6. The full U must be 3×3, so we need a third column orthonormal to the first two. Solve $\mathbf{u}_1 \cdot \mathbf{w} = \mathbf{u}_2 \cdot \mathbf{w} = 0$, that is

$$w_1 + w_2 + 2w_3 = 0, \qquad w_1 - w_2 = 0.$$

Worked example 1 (3/3): the full U

Step 6. The full U must be 3×3, so we need a third column orthonormal to the first two. Solve $\mathbf{u}_1 \cdot \mathbf{w} = \mathbf{u}_2 \cdot \mathbf{w} = 0$, that is

$$w_1 + w_2 + 2w_3 = 0, \qquad w_1 - w_2 = 0.$$

So $w_1 = w_2$ and $2w_1 = -2w_3$; taking $w_1 = 1$ gives $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, and normalising, $\mathbf{u}_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$.

Worked example 1 (3/3): the full U

Step 6. The full U must be 3×3, so we need a third column orthonormal to the first two. Solve $\mathbf{u}_1 \cdot \mathbf{w} = \mathbf{u}_2 \cdot \mathbf{w} = 0$, that is

$$w_1 + w_2 + 2w_3 = 0, \qquad w_1 - w_2 = 0.$$

So $w_1 = w_2$ and $2w_1 = -2w_3$; taking $w_1 = 1$ gives $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, and normalising, $\mathbf{u}_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$.

$$A = \underbrace{\begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ 2/\sqrt{6} & 0 & -1/\sqrt{3} \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}}_{\Sigma} \underbrace{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{V^{\mathsf{T}}}$$

Worked example 2: a zero singular value

Find an SVD of $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

Worked example 2: a zero singular value

Find an SVD of $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

Steps 1–4. $A^{\mathsf{T}}A = \left[\begin{smallmatrix} 2 & 2 \\ 2 & 2 \end{smallmatrix}\right]$ has eigenvalues 4 and 0, with eigenvectors $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$. So $s_1 = 2$, $s_2 = 0$, and $\mathbf{v}_1 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$, $\mathbf{v}_2 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$.

Worked example 2: a zero singular value

Find an SVD of $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

Steps 1–4. $A^{\mathsf{T}}A = \left[\begin{smallmatrix} 2 & 2 \\ 2 & 2 \end{smallmatrix}\right]$ has eigenvalues 4 and 0, with eigenvectors $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$. So $s_1 = 2$, $s_2 = 0$, and $\mathbf{v}_1 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$, $\mathbf{v}_2 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$.

Step 5. Only $s_1 > 0$, so $\mathbf{u}_1 = \frac{1}{2}A\mathbf{v}_1 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$.

Worked example 2: a zero singular value

Find an SVD of $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

Steps 1–4. $A^{\mathsf{T}}A = \left[\begin{smallmatrix} 2 & 2 \\ 2 & 2 \end{smallmatrix}\right]$ has eigenvalues 4 and 0, with eigenvectors $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$. So $s_1 = 2$, $s_2 = 0$, and $\mathbf{v}_1 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$, $\mathbf{v}_2 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$.

Step 5. Only $s_1 > 0$, so $\mathbf{u}_1 = \frac{1}{2}A\mathbf{v}_1 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$.

Step 6. We need one more column, orthonormal to $\mathbf{u}_1$. Take $\mathbf{u}_2 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$.

Worked example 2: a zero singular value

Find an SVD of $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

Steps 1–4. $A^{\mathsf{T}}A = \left[\begin{smallmatrix} 2 & 2 \\ 2 & 2 \end{smallmatrix}\right]$ has eigenvalues 4 and 0, with eigenvectors $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$. So $s_1 = 2$, $s_2 = 0$, and $\mathbf{v}_1 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$, $\mathbf{v}_2 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$.

Step 5. Only $s_1 > 0$, so $\mathbf{u}_1 = \frac{1}{2}A\mathbf{v}_1 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$.

Step 6. We need one more column, orthonormal to $\mathbf{u}_1$. Take $\mathbf{u}_2 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$.

$$A = \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}}_{\Sigma} \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{V^{\mathsf{T}}}$$

Check: $\Sigma V^{\mathsf{T}} = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 2 & 2 \\ 0 & 0 \end{smallmatrix}\right]$, and U times that is $\frac{1}{2}\left[\begin{smallmatrix} 2 & 2 \\ 2 & 2 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]$. **Yes.**

What the pieces mean

Let r be the number of non-zero singular values. Then:

RANK	Rank $A = r$
Im A	spanned by $\mathbf{u}_1, \ldots, \mathbf{u}_r$
Ker A	spanned by $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$
Coker A	spanned by $\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m$

What the pieces mean

Let r be the number of non-zero singular values. Then:

RANK	Rank $A = r$
Im A	spanned by $\mathbf{u}_1, \ldots, \mathbf{u}_r$
Ker A	spanned by $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$
Coker A	spanned by $\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m$

So a single factorisation hands us orthonormal bases of the fundamental subspaces all at once.

In example 2, $r = 1$: rank 1, with Ker A spanned by $\mathbf{v}_2 = \frac{1}{\sqrt{2}} \left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$, which A indeed kills.

SECTION 5

Least squares, revisited

The problem with the normal equations

In Lecture 18: $\hat{\mathbf{x}} = (A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}\mathbf{b}$.

- Forming $A^{\mathsf{T}}A$ squares the singular values, and so squares the ratio s_1/s_n, which governs how much small errors in the data are amplified.
- If A has dependent columns, $A^{\mathsf{T}}A$ is singular and the formula fails outright.
- Inverting an $n \times n$ matrix is expensive, and the result is then thrown away.

NOT EXAMINABLE

The problem with the normal equations

In Lecture 18: $\hat{\mathbf{x}} = (A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}\mathbf{b}$.

- Forming $A^{\mathsf{T}}A$ squares the singular values, and so squares the ratio s_1/s_n, which governs how much small errors in the data are amplified.
- If A has dependent columns, $A^{\mathsf{T}}A$ is singular and the formula fails outright.
- Inverting an $n \times n$ matrix is expensive, and the result is then thrown away.

KEY IDEA

The SVD sidesteps all three, because the only thing that ever needs inverting is a *diagonal* matrix.

NOT EXAMINABLE

The pseudoinverse

DEFINITION 20.10

Let $A = U\Sigma V^{\mathsf{T}}$ with r non-zero singular values. Define $\Sigma^{+} \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ by transposing Σ and replacing each non-zero s_i by $1/s_i$, leaving zeros as zeros. Then the **pseudoinverse** of A is

$$A^{+} \coloneqq V\Sigma^{+}U^{\mathsf{T}}.$$

NOT EXAMINABLE

DEFINITION 20.10

Let $A = U\Sigma V^{\mathsf{T}}$ with r non-zero singular values. Define $\Sigma^{+} \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ by transposing Σ and replacing each non-zero s_i by $1/s_i$, leaving zeros as zeros. Then the **pseudoinverse** of A is

$$A^{+} \coloneqq V\Sigma^{+}U^{\mathsf{T}}.$$

EXAMPLE 20.11

For $\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ we get $\Sigma^{+} = \begin{bmatrix} 1/\sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$.

NOT EXAMINABLE

The pseudoinverse

DEFINITION 20.10

Let $A = U\Sigma V^{\mathsf{T}}$ with r non-zero singular values. Define $\Sigma^{+} \in \mathrm{Mat}_{n\times m}(\mathbb{R})$ by transposing Σ and replacing each non-zero s_i by $1/s_i$, leaving zeros as zeros. Then the **pseudoinverse** of A is

$$A^{+} \coloneqq V\Sigma^{+}U^{\mathsf{T}}.$$

EXAMPLE 20.11

For $\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ we get $\Sigma^{+} = \begin{bmatrix} 1/\sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$.

Note what is *not* done: a zero singular value stays zero. We never divide by it, and that is exactly what keeps the construction defined when A is rank-deficient.

NOT EXAMINABLE

Why it solves least squares

THEOREM 20.12

$\hat{\mathbf{x}} = A^{+}\mathbf{b}$ *is a least-squares solution of* $A\mathbf{x} = \mathbf{b}$*. When* A *has independent columns it is the solution; otherwise it is the one of smallest length.*

NOT EXAMINABLE

Why it solves least squares

THEOREM 20.12

$\hat{\mathbf{x}} = A^{+}\mathbf{b}$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$. When A has independent columns it is the solution; otherwise it is the one of smallest length.

SKETCH.

Because U is orthogonal it preserves lengths, so

$$\|\mathbf{b} - A\mathbf{x}\| \;=\; \|U^{\mathsf{T}}\mathbf{b} - \Sigma V^{\mathsf{T}}\mathbf{x}\|.$$

Write $\mathbf{y} = V^{\mathsf{T}}\mathbf{x}$ and $\mathbf{c} = U^{\mathsf{T}}\mathbf{b}$. Since Σ is diagonal, $\|\mathbf{c} - \Sigma\mathbf{y}\|^2 = \sum_{i \le r}(c_i - s_i y_i)^2 + \sum_{i > r} c_i^2$. The first sum is killed by $y_i = c_i/s_i$; the second cannot be touched at all. □

NOT EXAMINABLE

Why it solves least squares

THEOREM 20.12

$\hat{\mathbf{x}} = A^{+}\mathbf{b}$ *is a least-squares solution of* $A\mathbf{x} = \mathbf{b}$*. When* A *has independent columns it is the solution; otherwise it is the one of smallest length.*

SKETCH.

Because U is orthogonal it preserves lengths, so

$$\|\mathbf{b} - A\mathbf{x}\| \;=\; \|U^{\mathsf{T}}\mathbf{b} - \Sigma V^{\mathsf{T}}\mathbf{x}\|.$$

Write $\mathbf{y} = V^{\mathsf{T}}\mathbf{x}$ and $\mathbf{c} = U^{\mathsf{T}}\mathbf{b}$. Since Σ is diagonal, $\|\mathbf{c} - \Sigma\mathbf{y}\|^2 = \sum_{i \le r}(c_i - s_i y_i)^2 + \sum_{i > r} c_i^2$. The first sum is killed by $y_i = c_i/s_i$; the second cannot be touched at all. □

The unreachable part of $\mathbf{b}$ is exactly the residual, and the y_i for $i > r$ are free, so setting them to zero gives the shortest $\hat{\mathbf{x}}$.

NOT EXAMINABLE

An example

Solve $A\mathbf{x} = \mathbf{b}$ in the least-squares sense, with $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.

NOT EXAMINABLE

An example

Solve $A\mathbf{x} = \mathbf{b}$ in the least-squares sense, with $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.

From worked example 1, $\Sigma^+ = \begin{bmatrix} 1/\sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, and

$$U^{\mathsf{T}}\mathbf{b} = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{6} \\ 0 \\ 2/\sqrt{3} \end{bmatrix}.$$

NOT EXAMINABLE

An example

Solve $A\mathbf{x} = \mathbf{b}$ in the least-squares sense, with $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.

From worked example 1, $\Sigma^+ = \begin{bmatrix} 1/\sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, and

$$U^{\mathsf{T}}\mathbf{b} = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{6} \\ 0 \\ 2/\sqrt{3} \end{bmatrix}.$$

$$\Sigma^+ U^{\mathsf{T}}\mathbf{b} = \begin{bmatrix} 2/\sqrt{18} \\ 0 \end{bmatrix}, \qquad \hat{\mathbf{x}} = V \begin{bmatrix} 2/\sqrt{18} \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2/\sqrt{18} \\ 0 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 1/3 \end{bmatrix}.$$

NOT EXAMINABLE

An example

Solve $A\mathbf{x} = \mathbf{b}$ in the least-squares sense, with $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.

From worked example 1, $\Sigma^+ = \begin{bmatrix} 1/\sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, and

$$U^{\mathsf{T}}\mathbf{b} = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{6} \\ 0 \\ 2/\sqrt{3} \end{bmatrix}.$$

$$\Sigma^+ U^{\mathsf{T}}\mathbf{b} = \begin{bmatrix} 2/\sqrt{18} \\ 0 \end{bmatrix}, \qquad \hat{\mathbf{x}} = V \begin{bmatrix} 2/\sqrt{18} \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2/\sqrt{18} \\ 0 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 1/3 \end{bmatrix}.$$

The same $\hat{\mathbf{x}}$ we found from the normal equations in Lecture 18, but with no matrix inverted except a diagonal one.

NOT EXAMINABLE

Where else this goes

- **Rank and conditioning**: the singular values say how close A is to being rank-deficient, which no determinant can.
- **Low-rank approximation**: keeping only the largest few s_i gives the best approximation of that rank, the basis of image compression and of latent-factor models.
- **Principal component analysis**: the $\mathbf{v}_i$ are the principal directions of a data cloud, in decreasing order of spread.
- **Numerical linear algebra**: the SVD is the standard tool for least squares precisely because it is stable.

NOT EXAMINABLE

Where else this goes

- **Rank and conditioning**: the singular values say how close A is to being rank-deficient, which no determinant can.
- **Low-rank approximation**: keeping only the largest few s_i gives the best approximation of that rank, the basis of image compression and of latent-factor models.
- **Principal component analysis**: the $\mathbf{v}_i$ are the principal directions of a data cloud, in decreasing order of spread.
- **Numerical linear algebra**: the SVD is the standard tool for least squares precisely because it is stable.

In every case one computes the SVD once and reads several answers off it. That economy is the reason it is worth the effort.

NOT EXAMINABLE

Summary of Lecture 20

SYMMETRIC	$A^{\mathsf{T}} = A$: real eigenvalues and an orthonormal eigenbasis.
KEY FACT	$A^{\mathsf{T}}A$ is symmetric and $n \times n$, and its eigenvalues satisfy $\lambda = \|A\mathbf{v}\|^2/\|\mathbf{v}\|^2 \geq 0$.
SINGULAR VALUES	$s_i = \sqrt{\lambda_i}$, ordered decreasing.
THEOREM	Every $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ factors as $A = U\Sigma V^{\mathsf{T}}$ with U, V orthogonal and Σ the singular values on the diagonal.
GEOMETRY	Read right to left: rotate, stretch each axis by s_i, rotate.
RECIPE	$A^{\mathsf{T}}A \to$ eigenvalues $\to s_i = \sqrt{\lambda_i} \to$ orthonormal eigenvectors give $V \to \mathbf{u}_i = A\mathbf{v}_i/s_i \to$ extend to U if the full form is wanted.
READS OFF	$\mathrm{Rank}\, A = r$; $\mathrm{Im}\, A$ from $\mathbf{u}_1, \ldots, \mathbf{u}_r$, $\mathrm{Ker}\, A$ from $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$.

Before next time

AT HOME

1. Find an SVD of $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$, and of $B = \begin{bmatrix} 0 & 2 \\ 3 & 0 \end{bmatrix}$.
2. Find an SVD of $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, and read off $\operatorname{Rank} A$, a basis of $\operatorname{Im} A$ and a basis of $\operatorname{Ker} A$.
3. Find a reduced SVD of $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 2 & -1 \end{bmatrix}$.
4. Show that if $A^2 = I$, then every eigenvalue of A is 1 or -1.

END OF LECTURE 20

ilia.pirashvili@universityofgalway.ac.ie

Solution: question 1

FOR A

$A^{\mathsf{T}}A = \left[\begin{smallmatrix} 4 & 0 \\ 0 & 9 \end{smallmatrix}\right]$, eigenvalues $9, 4$. So $s_1 = 3$, $s_2 = 2$, with $\mathbf{v}_1 = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$, $\mathbf{v}_2 = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$. Then $\mathbf{u}_1 = \frac{1}{3} A \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$ and $\mathbf{u}_2 = \frac{1}{2} A \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$:

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Solution: question 1

FOR A

$A^{\mathsf{T}}A = \left[\begin{smallmatrix} 4 & 0 \\ 0 & 9 \end{smallmatrix}\right]$, eigenvalues $9, 4$. So $s_1 = 3$, $s_2 = 2$, with $\mathbf{v}_1 = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$, $\mathbf{v}_2 = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$. Then $\mathbf{u}_1 = \frac{1}{3}A\left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$ and $\mathbf{u}_2 = \frac{1}{2}A\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$:

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

FOR B

$B^{\mathsf{T}}B = \left[\begin{smallmatrix} 9 & 0 \\ 0 & 4 \end{smallmatrix}\right]$, so $s_1 = 3$, $s_2 = 2$ with $\mathbf{v}_1 = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$, $\mathbf{v}_2 = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$, giving $\mathbf{u}_1 = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$, $\mathbf{u}_2 = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$.

Solution: question 2

SOLUTION

$A^{\mathsf{T}}A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, eigenvalues 2 and 0, so $s_1 = \sqrt{2}$ and $s_2 = 0$, with $\mathbf{v}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

$\mathbf{u}_1 = \frac{1}{\sqrt{2}} A\mathbf{v}_1 = \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}}\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, and $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ completes $U = I_2$.

Solution: question 2

SOLUTION

$A^{\mathsf{T}}A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, eigenvalues 2 and 0, so $s_1 = \sqrt{2}$ and $s_2 = 0$, with $\mathbf{v}_1 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$, $\mathbf{v}_2 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$.

$\mathbf{u}_1 = \frac{1}{\sqrt{2}} A\mathbf{v}_1 = \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 2 \\ 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$, and $\mathbf{u}_2 = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$ completes $U = I_2$.

So $r = 1$: $\operatorname{Rank} A = 1$, $\operatorname{Im} A = \operatorname{Span}\{\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]\}$, and $\operatorname{Ker} A = \operatorname{Span}\{\left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]\}$.

All three read off the decomposition, with no further computation.

Solution: question 3

SOLUTION

$A^{\mathsf{T}}A = \begin{bmatrix} 9 & 2 \\ 2 & 6 \end{bmatrix}$, and $\lambda^2 - 15\lambda + 50 = (\lambda - 10)(\lambda - 5)$, so $s_1 = \sqrt{10}$, $s_2 = \sqrt{5}$. The kernels give $\mathbf{v}_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ -2 \end{bmatrix}$.

$\mathbf{u}_1 = \frac{1}{\sqrt{10}} A\mathbf{v}_1 = \frac{1}{\sqrt{50}} \begin{bmatrix} 4 \\ 5 \\ 3 \end{bmatrix}$, and $\mathbf{u}_2 = \frac{1}{\sqrt{5}} A\mathbf{v}_2 = \frac{1}{5} \begin{bmatrix} -3 \\ 0 \\ 4 \end{bmatrix}$.

$$A = \begin{bmatrix} 4/\sqrt{50} & -3/5 \\ 5/\sqrt{50} & 0 \\ 3/\sqrt{50} & 4/5 \end{bmatrix} \begin{bmatrix} \sqrt{10} & 0 \\ 0 & \sqrt{5} \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}.$$

Solution: question 4

SOLUTION

Let $A\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} \neq \mathbf{0}$. Applying A again,

$$\mathbf{v} = I\mathbf{v} = A^2\mathbf{v} = A(\lambda\mathbf{v}) = \lambda A\mathbf{v} = \lambda^2\mathbf{v},$$

so $(\lambda^2 - 1)\mathbf{v} = \mathbf{0}$. As $\mathbf{v} \neq \mathbf{0}$, we get $\lambda^2 = 1$, that is $\lambda = \pm 1$.