

Reduction II: What the SVD Actually Does

A matrix, a circle, and three programs

James Pustejovsky

Brandeis University
Department of Computer Science

CS 135: Computational Semantics

Fall 2026

Where We Left Off

Session 8 used the SVD as a **tool**. We wrote $M = U\Sigma V^\top$, read off the singular values, dropped the small ones, and the word vectors got better.

We never asked what the factorization *is*.

Today

One 2×2 matrix, one circle. We apply the matrix directly, then apply the three factors one at a time, and check that the two agree. By the end, U , Σ and $V^\top$ each have a geometric job description.

This also sets up the geometric algebra thread. Two of the three factors turn out to be *rotors*, which is the object we build the second half of the course on.

Factoring Exposes Structure

Writing $12 = 3 \times 4$ does not change the number. It exposes what the number is built from, and it lets you answer questions you could not answer before: is it divisible by 6, how many divisors does it have.

$12 = 3 \times 4$	exposes the arithmetic factors
$A = U\Sigma V^\top$	exposes the <i>geometric</i> factors

A matrix is a single sudden transformation: numbers in, numbers out. The factorization separates that one action into three actions of three different kinds, applied in order.

The claim to test today: running the three in sequence gives exactly what A gives in one stroke.

The Running Matrix

$$A = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}$$

Small enough to compute by hand, and not symmetric, which matters: for a symmetric matrix the first and last factors coincide and the interesting asymmetry disappears.

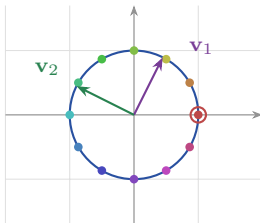
$$\det A = 4 \quad \|A\|_F = \sqrt{17} \approx 4.123 \quad A \text{ is invertible}$$

We will keep this matrix for the whole session, including the degenerate cases at the end, so that every picture is comparable to every other picture.

The Input: the Unit Circle

The unit circle is the set of all vectors of length one:

$$S^1 = \{ \mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| = 1 \} = \{ (\cos t, \sin t) : 0 \leq t < 2\pi \}$$



The twelve colored dots are placed every 30° . They carry no mathematics; they are there so that a transformation which does not change the *shape* is still visible. The two arrows are marked now and explained later.

Applying A Means Applying It Pointwise

There is no special "circle" operation. A acts on vectors, and the image of the circle is just the set of images of its points:

$$A \cdot S^1 = \{ A\mathbf{x} : \|\mathbf{x}\| = 1 \}$$

For $\mathbf{x} = (x_1, x_2)$:

$$A\mathbf{x} = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 + 3x_2 \\ 2x_2 \end{pmatrix}$$

Every point moves independently, by the same rule. Nothing in that rule mentions circles, ellipses, axes or angles. Whatever structure appears in the output was already implicit in the four numbers.

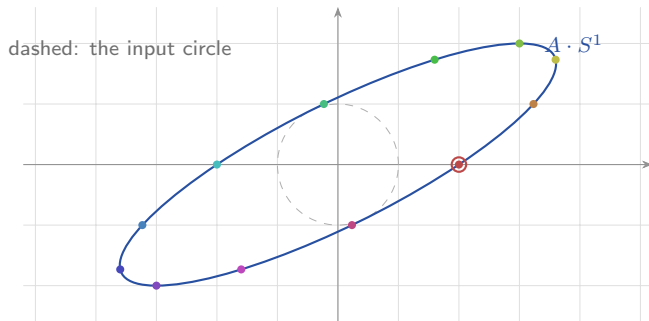
Four Points, By Hand

t	$\mathbf{x} = (\cos t, \sin t)$	$A\mathbf{x} = (2x_1 + 3x_2, 2x_2)$	$\ A\mathbf{x}\ $
0°	$(1, 0)$	$(2, 0)$	2.000
90°	$(0, 1)$	$(3, 2)$	3.606
63.43°	$(0.447, 0.894)$	$(3.578, 1.789)$	4.000
153.43°	$(-0.894, 0.447)$	$(-0.447, 0.894)$	1.000

Check the third row: $2(0.447) + 3(0.894) = 0.894 + 2.683 = 3.578$, and $2(0.894) = 1.789$.

Unit vectors go in. Vectors of four different lengths come out. The transformation stretches some directions more than others, and that is the entire subject of this lecture.

The Whole Circle at Once



The fact this lecture explains

The image of the unit circle under *any* 2×2 matrix is an ellipse centered at the origin. Not sometimes. Always.

Why It Has to Be an Ellipse

A quick argument, before the machinery. A linear map sends the circle to *some* closed curve, and the curve satisfies

$$\mathbf{y} = A\mathbf{x}, \quad \|\mathbf{x}\| = 1 \quad \implies \quad \mathbf{x} = A^{-1}\mathbf{y}, \quad \|A^{-1}\mathbf{y}\| = 1$$

Squaring the second condition gives

$$\mathbf{y}^\top (A^{-1})^\top A^{-1} \mathbf{y} = 1$$

which is $\mathbf{y}^\top Q \mathbf{y} = 1$ for a symmetric positive definite Q . That is the defining equation of an ellipse.

The argument assumes A is invertible. When it is not, the ellipse degenerates to a segment or a point, which is the last section of today.

First Guess: Are the Columns the Axes?

The columns of A are the images of the standard basis vectors:

$$A\mathbf{e}_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad A\mathbf{e}_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

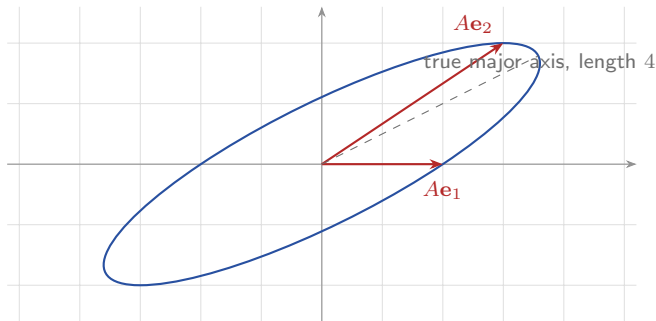
Both are on the ellipse, since $\mathbf{e}_1$ and $\mathbf{e}_2$ are on the circle. So it is tempting to read them as the axes.

$$\|A\mathbf{e}_1\| = 2 \quad \|A\mathbf{e}_2\| = \sqrt{13} \approx 3.606$$

$$A\mathbf{e}_1 \cdot A\mathbf{e}_2 = (2)(3) + (0)(2) = 6$$

They are not perpendicular, so they cannot be the two axes of an ellipse. And neither length is right: the true major semi-axis is 4, which is longer than both.

Seeing the Near Miss



Ae_2 points near the long direction but overshoots the angle and undershoots the length. The standard basis is simply not aligned with anything A cares about.

So the question becomes: which input directions do land on the axes? That question is what the SVD answers.

What We Can and Cannot Read Off

Quantity	From A directly?	Value
area scale factor	yes, $ \det A $	4
total size	yes, $\ A\ _F$	$\sqrt{17} \approx 4.123$
length of major axis	no	4
length of minor axis	no	1
direction of the axes	no	$26.57^\circ, 116.57^\circ$
which inputs map there	no	$63.43^\circ, 153.43^\circ$

Everything in the lower block requires factoring A . Note the top block is consistent with the bottom: $\det A = 4 = 4 \times 1 = \sigma_1 \sigma_2$, and $\|A\|_F^2 = 17 = 16 + 1 = \sigma_1^2 + \sigma_2^2$.

The Factorization

Theorem

Every real matrix $A \in \mathbb{R}^{m \times n}$ can be written

$$A = U \Sigma V^\top$$

with U and V orthogonal and Σ diagonal with non-negative entries $\sigma_1 \geq \sigma_2 \geq \dots \geq 0$.

Factor	Kind of matrix	Freedom it has
$V^\top$	orthogonal	rotation or reflection only
Σ	diagonal, non-negative	independent scaling per axis
U	orthogonal	rotation or reflection only

Two of the three factors are rigid. All of the shape change in A has been forced into the middle one.

Our Matrix, Factored

$$\underbrace{\begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}}_A = \underbrace{\frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}}_U \underbrace{\begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}}_{\Sigma} \underbrace{\frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}}_{V^T}$$

In decimals:

$$U \approx \begin{pmatrix} 0.894 & -0.447 \\ 0.447 & 0.894 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \quad V^T \approx \begin{pmatrix} 0.447 & 0.894 \\ -0.894 & 0.447 \end{pmatrix}$$

Everything is exact: the singular values are 4 and 1, and every entry of U and V is $\pm 1/\sqrt{5}$ or $\pm 2/\sqrt{5}$. This matrix was chosen so that nothing rounds.

Checking the Factorization

Multiply the right-hand side back out. First ΣV^T :

$$\begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} \cdot \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 4 & 8 \\ -2 & 1 \end{pmatrix}$$

Then U times that, with the two $\sqrt{5}$ factors combining to 5:

$$\frac{1}{5} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 4 & 8 \\ -2 & 1 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 8+2 & 16-1 \\ 4-4 & 8+2 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 10 & 15 \\ 0 & 10 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}$$

The factorization is an identity, not an approximation. Nothing has been lost or rounded away.

Three Programs

Read $A\mathbf{x} = U(\Sigma(V^\top \mathbf{x}))$ right to left, as a pipeline. Each factor is a program with an input shape and an output shape.

	Program	Signature	Changes the shape?
1	$V^\top$	circle $\rightarrow$ circle	no
2	Σ	circle $\rightarrow$ axis-aligned ellipse	yes
3	U	ellipse $\rightarrow$ ellipse	no

Exactly one of the three changes the shape. The other two only decide orientation.

Program 1: $V^\top$, Computationally

$$V^\top = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \approx \begin{pmatrix} 0.447 & 0.894 \\ -0.894 & 0.447 \end{pmatrix}$$

Two checks that this is a rotation:

- Columns are unit length: $\frac{1}{5}(1+4) = 1$. ✓
- Columns are perpendicular: $\frac{1}{5}((1)(2) + (-2)(1)) = 0$. ✓
- $\det V^\top = \frac{1}{5}(1+4) = +1$, so it is a rotation and not a reflection. ✓

Matching against $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ gives $\cos \theta = 0.447$, $\sin \theta = -0.894$, so

$$\theta = -63.43^\circ$$

-63.43° is $-\arctan 2$. The rotation is clockwise.

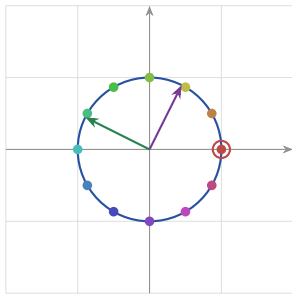
Program 1: the Arithmetic

$\mathbf{x}$	$V^T \mathbf{x}$	
$(1, 0)$	$(0.447, -0.894)$	
$(0, 1)$	$(0.894, 0.447)$	
$\mathbf{v}_1 = (0.447, 0.894)$	$(1, 0) = \mathbf{e}_1$	lands on the x axis
$\mathbf{v}_2 = (-0.894, 0.447)$	$(0, 1) = \mathbf{e}_2$	lands on the y axis

Row three, worked out: $(0.447)(0.447) + (0.894)(0.894) = 0.2 + 0.8 = 1$, and $(-0.894)(0.447) + (0.447)(0.894) = -0.4 + 0.4 = 0$.

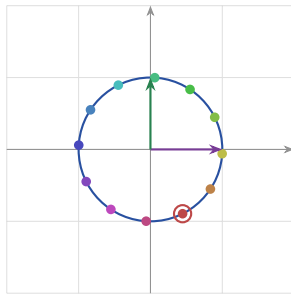
Every output has length one, as it must. The last two rows are the whole point of this program.

Program 1: the Picture



in

V^T



out

The outline is identical. Only the dots moved, clockwise by 63.43° . The purple and green arrows have come to rest exactly on the axes.

Program 1: What It Is Doing

Computationally	Geometrically
Rewrites each input vector in the basis $\{\mathbf{v}_1, \mathbf{v}_2\}$. Row i of $V^\top$ is $\mathbf{v}_i^\top$, so entry i of the output is $\mathbf{v}_i \cdot \mathbf{x}$: the component of $\mathbf{x}$ along $\mathbf{v}_i$.	Turns the plane so that the direction about to be stretched most lies along x , and the direction about to be stretched least lies along y . Lengths and angles are preserved, so the circle is untouched as a set.

The same sentence twice

"Express $\mathbf{x}$ in the $\mathbf{v}$ coordinate system" and "rotate the space until the $\mathbf{v}$ directions are the axes" describe one operation. A change of basis to an orthonormal basis *is* a rotation.

Program 2: Σ , Computationally

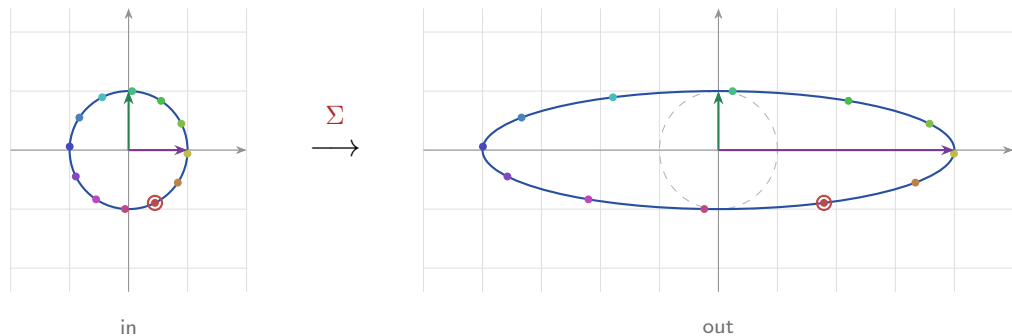
$$\Sigma = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \quad \Sigma \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 4y_1 \\ 1y_2 \end{pmatrix}$$

There is no mixing. Coordinate one is multiplied by 4 and coordinate two by 1, and neither consults the other.

$\mathbf{y}$	$\Sigma \mathbf{y}$
$(0.447, -0.894)$	$(1.789, -0.894)$
$(0.894, 0.447)$	$(3.578, 0.447)$
$(1, 0)$	$(4, 0)$
$(0, 1)$	$(0, 1)$

The second singular value is 1 here, so the y coordinate happens to be left alone. That is a property of this matrix, not of Σ in general.

Program 2: the Picture



This is the step where the circle stops being a circle. The semi-axes of the result are $\sigma_1 = 4$ horizontally and $\sigma_2 = 1$ vertically, read straight off the diagonal.

Program 2: What It Is Doing

Computationally	Geometrically
Scales each coordinate independently. In the $\mathbf{v}$ coordinate system installed by program 1, this is a per-coordinate multiply: no additions, no cross terms.	Stretches the plane along the two axes by σ_1 and σ_2 . The unit circle becomes an axis-aligned ellipse whose semi-axis lengths <i>are</i> the singular values.

The only factor that is not orthogonal

Σ does not preserve lengths, does not preserve angles, and is the only one of the three that can be non-invertible. All three facts are the same fact.

It is also the only factor that can destroy information, which is why truncation is always an operation on Σ .

Program 3: U , Computationally

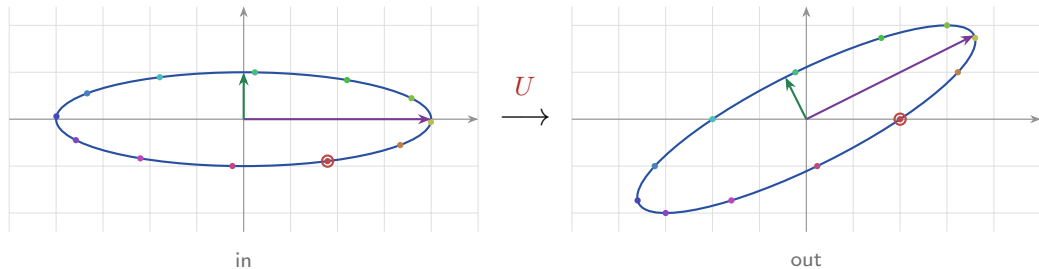
$$U = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \approx \begin{pmatrix} 0.894 & -0.447 \\ 0.447 & 0.894 \end{pmatrix}$$

Same three checks: columns of unit length, columns perpendicular, $\det U = \frac{1}{5}(4 + 1) = +1$.
Another rotation.

Matching the rotation template gives $\cos \theta = 0.894$, $\sin \theta = 0.447$, so $\theta = +26.57^\circ$, which is $\arctan \frac{1}{2}$, counterclockwise.

$\mathbf{z}$	$U\mathbf{z}$
(1.789, -0.894)	(2, 0)
(3.578, 0.447)	(3, 2)
(4, 0)	(3.578, 1.789)
(0, 1)	(-0.447, 0.894)

Program 3: the Picture



The axis lengths were fixed by Σ and are untouched here. This program decides only which way the finished ellipse points.

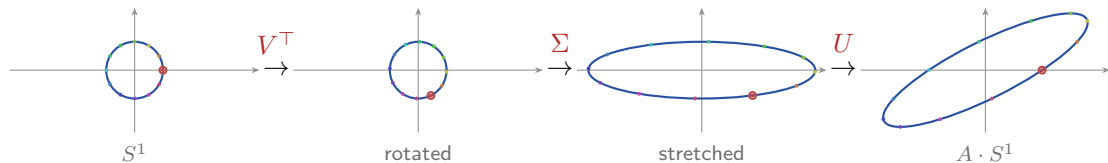
Program 3: What It Is Doing

Computationally	Geometrically
Reads the stretched coordinates as coefficients in the basis $\{\mathbf{u}_1, \mathbf{u}_2\}$. Column i of U is $\mathbf{u}_i$, so $U\mathbf{z} = z_1\mathbf{u}_1 + z_2\mathbf{u}_2$.	Rotates the axis-aligned ellipse into the orientation A actually produces, carrying the x axis to $\mathbf{u}_1$ and the y axis to $\mathbf{u}_2$.

Program 1 took the input out of the standard basis. Program 3 puts the output back into it. They are the two ends of a sandwich, and Σ is the filling.

Note $V^\top$ reads components off with dot products, while U builds a vector up from coefficients. That is the difference between a matrix acting by rows and by columns.

The Whole Pipeline



Rotate, stretch, rotate. The first and last panels are the two sides of the identity $A = U\Sigma V^T$.

One Point, All the Way Through

Take $\mathbf{x} = \mathbf{e}_1 = (1, 0)$, the red dot.

Stage	Vector	Length	What changed
input	$(1, 0)$	1.000	–
after $V^\top$	$(0.447, -0.894)$	1.000	direction only
after Σ	$(1.789, -0.894)$	2.000	first coordinate $\times 4$
after U	$(2, 0)$	2.000	direction only
$A\mathbf{x}$ directly	$(2, 0)$	2.000	same

The vector leaves at the same place it would have gone in one step, having visited two positions no one would have predicted from A alone.

A Second Point

Take $\mathbf{x} = \mathbf{e}_2 = (0, 1)$.

Stage	Vector	Length
input	$(0, 1)$	1.000
after $V^\top$	$(0.894, 0.447)$	1.000
after Σ	$(3.578, 0.447)$	3.606
after U	$(3, 2)$	3.606
$A\mathbf{x}$ directly	$(3, 2)$	3.606

$3.606 = \sqrt{13}$, the length we computed on the columns slide. The stretch factor for this particular input is 3.606, which sits between $\sigma_2 = 1$ and $\sigma_1 = 4$, as every stretch factor must.

The Point That Explains Everything

Now take $\mathbf{x} = \mathbf{v}_1 = (0.447, 0.894)$.

Stage	Vector	Note
input	$\mathbf{v}_1 = (0.447, 0.894)$	on the circle
after $V^\top$	$(1, 0) = \mathbf{e}_1$	exactly the first axis
after Σ	$(4, 0) = \sigma_1 \mathbf{e}_1$	scaled by σ_1 and nothing else
after U	$(3.578, 1.789) = \sigma_1 \mathbf{u}_1$	the end of the major axis

The fundamental relation

$$A\mathbf{v}_i = \sigma_i \mathbf{u}_i$$

$\mathbf{v}_i$ is the input direction; $\mathbf{u}_i$ is where it comes out; σ_i is how much longer it got. The middle program is a plain multiplication *because* program 1 put the vector on an axis first.

Reading the Relation Off the Original Matrix

The relation can be checked against A with no factorization in sight:

$$\begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix} \cdot \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2+6 \\ 0+4 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 8 \\ 4 \end{pmatrix} = 4 \cdot \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

and for the second pair,

$$\begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix} \cdot \frac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} -4+3 \\ 0+2 \end{pmatrix} = 1 \cdot \frac{1}{\sqrt{5}} \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Two special input directions come out pointing along perpendicular output directions, scaled and not rotated relative to each other. Finding those directions is what computing an SVD means.

Answering the Question From Earlier

We asked: which input directions land on the axes of the ellipse?

	Input direction	Output
major axis	$\mathbf{v}_1$ at 63.43°	$\sigma_1 \mathbf{u}_1$, length 4, at 26.57°
minor axis	$\mathbf{v}_2$ at 153.43°	$\sigma_2 \mathbf{u}_2$, length 1, at 116.57°

The two input directions are 90° apart. So are the two output directions. **But the pair has been turned:** the inputs sit at 63.43° and the outputs at 26.57° .

A perpendicular pair goes in and a perpendicular pair comes out, but it is a *different* perpendicular pair. That gap between $V^\top$ and U is exactly why a single rotation could never have described A .

What Survives Each Stage

Preserved?	V^T	Σ	U
lengths	yes	no	yes
angles	yes	no	yes
area	yes	no, $\times 4$	yes
orientation	yes	yes	yes
information	all	all, here	all

The area column multiplies out as $1 \times 4 \times 1 = 4 = \det A$. The determinant of a product is the product of the determinants, and the two rotations contribute 1 each.

The last row says "here" because $\sigma_2 = 1 \neq 0$. The next slides change that.

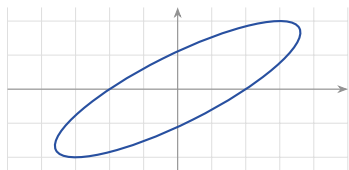
Where the Numbers Live

Question about A	Answered by	Value here
How much does it stretch, at most?	σ_1	4
How much, at least?	σ_2	1
Which input achieves the most?	$\mathbf{v}_1$, column 1 of V	63.43°
Where does that input land?	$\mathbf{u}_1$, column 1 of U	26.57°
By how much does area grow?	$\sigma_1\sigma_2$	4
Is it invertible?	is $\sigma_2 > 0$?	yes
How close to singular?	σ_1/σ_2	4

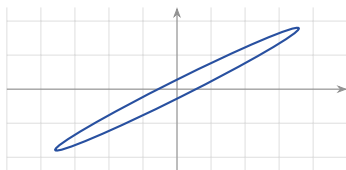
Every one of these is a question about A . None of them can be read off the four entries of A , and all of them are immediate once it is factored. That is what "exposing the structure" means in practice.

Degenerate Case: a Small σ_2

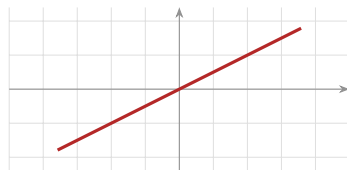
Keep U and V and shrink the second singular value.



$$\sigma_2 = 1$$



$$\sigma_2 = 0.25$$



$$\sigma_2 = 0$$

As $\sigma_2 \rightarrow 0$ the ellipse flattens onto the line through $\mathbf{u}_1$. At $\sigma_2 = 0$ the matrix has rank 1 and the whole plane is crushed onto that single line.

The map stops being invertible exactly when σ_2 reaches zero, and the ratio σ_1/σ_2 measures how close it is to that. This is the condition number.

Degenerate Case: a Reflection

Flip one sign: $A' = \begin{pmatrix} 2 & 3 \\ 0 & -2 \end{pmatrix}$, so $\det A' = -4$.

The singular values are unchanged: $\sigma_1 = 4$, $\sigma_2 = 1$. **The output ellipse is the same ellipse, the same shape in the same place.**

What differs is the traversal. Under A the colored dots run counterclockwise; under A' they run clockwise.

	A	A'
$\det$	+4	-4
σ_1, σ_2	4, 1	4, 1
$\det U$	+1, a rotation	-1, a reflection

Σ is non-negative by convention and V is taken to be a rotation, so when the map reverses orientation the reflection has to be carried by U . It cannot be hidden anywhere else.

Singular Values Cannot See Orientation

This is worth stating on its own, because it will matter later.

σ_1 and σ_2 describe the *shape* of the image and nothing else. Two matrices with the same singular values produce the same ellipse, and may still differ by a reflection.

Orientation lives in the sign of $\det$, and $\det A = \det U \cdot \det \Sigma \cdot \det V^T$ with the two rotations contributing ± 1 .

In the semantic setting this is the reason a reduced space can be right about similarity and wrong about direction. Cosine sees Σ ; it does not see the sign.

Truncation, in This Picture

Session 8 truncated by keeping the k largest singular values. Geometrically that is:

Algebra	Geometry
set $\sigma_{k+1} = \dots = \sigma_r = 0$	flatten the ellipsoid along those axes
keep $\sigma_1, \dots, \sigma_k$	keep the k longest axes at full length
$\ M - M_k\ _F^2 = \sum_{i>k} \sigma_i^2$	the total length that was flattened away

Eckart–Young says: of all the ways to flatten an ellipsoid down to k dimensions, flattening the *shortest* axes is the one that moves the surface least. Stated that way it is close to obvious. The content of the theorem is that no cleverer, non-axis-aligned flattening does better.

And the loss is always inflicted on Σ , because Σ is the only factor that has any information to lose.

Two of the Three Are Rotors

An orthogonal matrix with determinant $+1$ is a rotation. In geometric algebra a rotation is carried by a **rotor**

$$R = \exp\left(-\frac{1}{2}\theta \mathbf{e}_{12}\right) = \cos \frac{\theta}{2} - \sin \frac{\theta}{2} \mathbf{e}_{12}$$

applied by the sandwich $\mathbf{x} \mapsto R \mathbf{x} \tilde{R}$.

For our two factors:

	angle θ	half-angle in the rotor
$V^\top$	-63.43°	-31.72°
U	$+26.57^\circ$	$+13.28^\circ$

The half-angle is not a convenience of notation. It is forced by the two-sided sandwich, and it is the first real structural difference between rotors and rotation matrices.

Versors, and What Σ Is Not

Reflections and rotations together are the **versors**: the transformations a geometric algebra implements by sandwiching. Every orthogonal map is a versor.

Factor	Orthogonal?	A versor?
$V^\top$	yes	yes, a rotor
Σ	no	no
U	yes	yes, a rotor or a reflection

The factorization, restated

Every linear map is a versor, then a stretch along a frame, then another versor. Everything that is not length-preserving has been collected into the middle, and that middle is where reduction does its work.

Why This Matters for the Rest of the Course

When we move to geometric algebra we stop writing transformations as arrays of numbers and start writing them as elements of the algebra itself.

Matrix language	Geometric algebra language
orthogonal matrix, $\det = +1$	rotor R , applied as $Rx\tilde{R}$
composing two rotations	multiplying two rotors
the plane of rotation	the bivector e_{12}
the angle	the magnitude of the bivector

The advantage is not notational economy. It is that the plane of rotation becomes a first-class object you can name, add and compose, rather than something recovered after the fact from a matrix.

In $\mathbb{R}^2$ there is only one plane, so the benefit is invisible. It appears the moment we go to three dimensions and beyond.

Summary

Factor	Computationally	Geometrically
$V^\top$	take dot products with the $\mathbf{v}_i$	rotate so the extreme directions become the axes
Σ	multiply coordinate i by σ_i	stretch the circle into an axis-aligned ellipse
U	build the output from the $\mathbf{u}_i$	rotate the ellipse to its final orientation

- Every matrix maps the unit circle to an ellipse.
- σ_1, σ_2 are the semi-axis lengths; $\mathbf{u}_1, \mathbf{u}_2$ their directions; $\mathbf{v}_1, \mathbf{v}_2$ the inputs that get there.
- $A\mathbf{v}_i = \sigma_i\mathbf{u}_i$ is the whole factorization in one line.
- The columns of A are on the ellipse but are not its axes.
- Only Σ changes shape, and only Σ can lose information.

Exercises

- 1 For $B = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$, write down U , Σ , $V^\top$ without computing anything. What are $\mathbf{v}_1$ and $\mathbf{u}_1$?
- 2 For $C = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix}$, the image of the unit circle is a circle of radius 2. What does that say about σ_1 and σ_2 ? Are $\mathbf{v}_1$ and $\mathbf{u}_1$ determined?
- 3 Our A has $\|A\mathbf{e}_2\| = \sqrt{13} \approx 3.606$, between $\sigma_2 = 1$ and $\sigma_1 = 4$. Explain why no unit vector can have $\|A\mathbf{x}\|$ outside $[\sigma_2, \sigma_1]$.
- 4 Take A and set σ_2 to 0, keeping U and V . Write out the resulting rank-1 matrix as $\sigma_1 \mathbf{u}_1 \mathbf{v}_1^\top$ and check that its columns are multiples of $\mathbf{u}_1$.