Edit any entry (range −9 to 9). Each cell's fill height is proportional to a_{ij}^2. Teal = positive, amber = negative.



The Frobenius Norm

Robin Whitty

Maths Study Group 06/11/26

SUM OF SQUARED ENTRIES — A_{ij}^2

0.00 + 1.00 + 4.00 + 1.00 + 0.00 + 9.00 + 4.00 + 9.00 + 0.00 = 28.0000

FROBENIUS NORM

5.2915

$\|A\|_F = \sqrt{(28.0000)}$

FROBENIUS NORM SQUARED

28.0000

$\|A\|_F^2$

TR(A^TA) — TRACE FORMULA

28.0000

equals $\|A\|_F^2$ ✓

SINGULAR VALUES Σ_i

3.742 3.742 0.000

σ_1 σ_2 σ_3

$\Sigma\sigma_i^2 = 28.0000 = \|A\|_F^2$ ✓

Matrix norms

A matrix norm is a mapping from $m \times n$ matrices to the positive reals.

The idea is to measure how 'big' the matrix is or, equivalently, how big is the distance between two matrices.

Write $\|A\|$ for (some version of) the norm of matrix A. We require:

1. $\|A\| \geq 0$, with equality only for the all-zeros matrix;
2. for scalar α, $\|\alpha A\| = |\alpha| \|A\|$;
3. $\|A + B\| \leq \|A\| + \|B\|$ (triangle inequality).

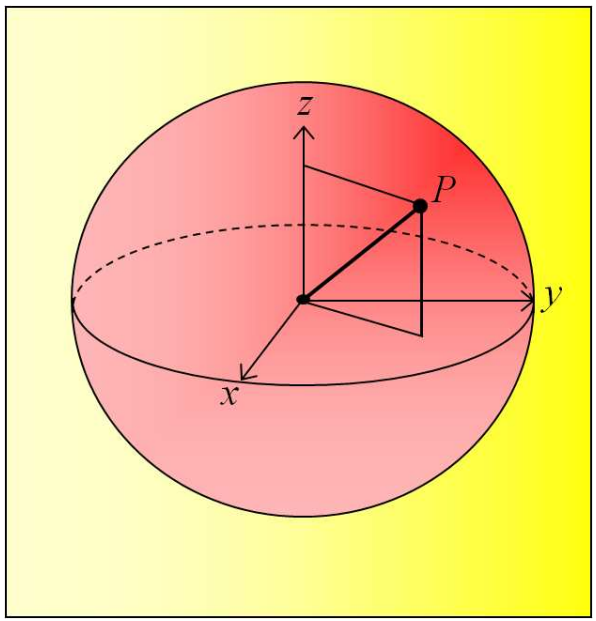
Most matrix norms are really just vector norms: we think of the $m \times n$ matrix A as a concatenation of its m rows to form a vector of length mn. However, a matrix norm can summarise the analysis of a matrix, notably relating to its **singular value decomposition**.

The Frobenius norm

Let A be an $m \times n$ matrix over the reals. The Frobenius norm of A is defined to be

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2}.$$

As a calculation this is really just the Euclidean distance norm on a vector. Thus in the following picture:



if point P is specified as a column vector of coordinates $P = (a, b, c)^T$ then the distance of P from the origin is $\|P\|_F$ (treating P as a 3×1 matrix).

The matrix inner product

Recall that the **trace** of a square matrix A, denoted $\mathrm{Tr}(A)$, is the sum of its diagonal elements.

Take the space of $m \times n$ matrices. We can make this an inner product space by defining the inner (scalar) product of A and B as $\langle A, B \rangle = \mathrm{Tr}(A^T B)$.

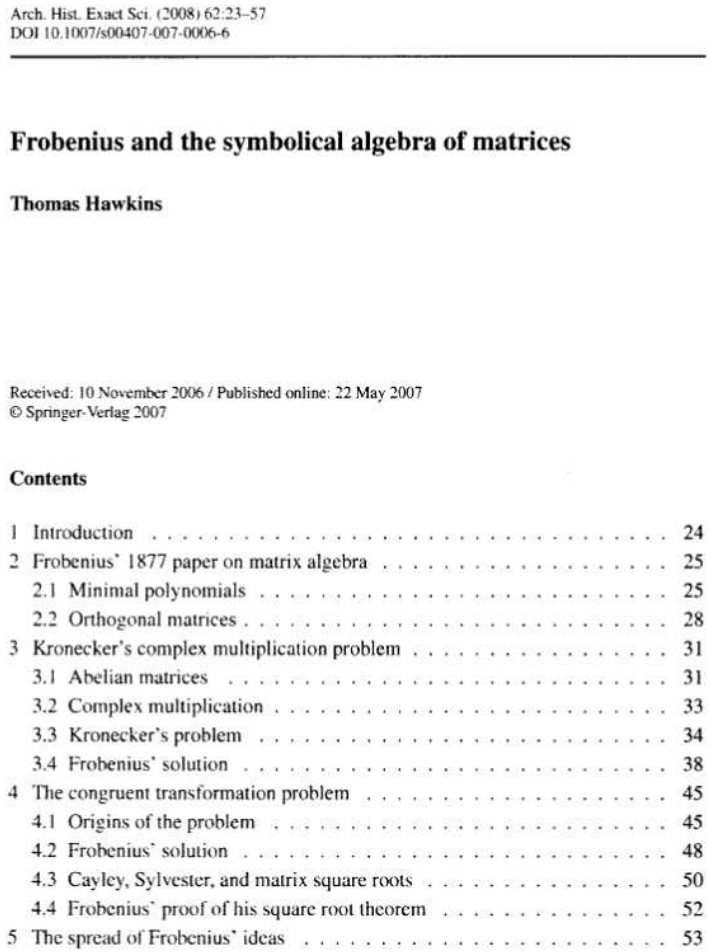
E.g. $A = \begin{pmatrix} 2 & 0 & 3 \\ 3 & 1 & -2 \end{pmatrix}, B = \begin{pmatrix} 5 & 3 & 2 \\ -2 & -2 & 3 \end{pmatrix}$

$$\begin{aligned} \langle A, B \rangle &= \mathrm{Tr}\left[\begin{pmatrix} 2 & 3 \\ 0 & 1 \\ 3 & -2 \end{pmatrix} \times \begin{pmatrix} 5 & 3 & 2 \\ -2 & -2 & 3 \end{pmatrix} \right] \\ &= \mathrm{Tr} \begin{pmatrix} 4 & 0 & 13 \\ -2 & -2 & 3 \\ 19 & 13 & 0 \end{pmatrix} = 2. \end{aligned}$$

Frobenius norm as a matrix trace

The inner product $\langle A, B \rangle$ is often called the Frobenius inner product (I think this came first and gave its name to the norm. Frobenius was among those making linear algebra rigorous in the second half of the 19th century).

We can transfer the trace calculation to $\|A\|_F$ (which is actually our main interest).



Arch. Hist. Exact Sci. (2008) 62:23–57
DOI 10.1007/s00407-007-0006-6

Frobenius and the symbolical algebra of matrices

Thomas Hawkins

Received: 10 November 2006 / Published online: 22 May 2007
© Springer-Verlag 2007

Contents

1 Introduction 24
2 Frobenius' 1877 paper on matrix algebra 25
2.1 Minimal polynomials 25
2.2 Orthogonal matrices 28
3 Kronecker's complex multiplication problem 31
3.1 Abelian matrices 31
3.2 Complex multiplication 33
3.3 Kronecker's problem 34
3.4 Frobenius' solution 38
4 The congruent transformation problem 45
4.1 Origins of the problem 45
4.2 Frobenius' solution 48
4.3 Cayley, Sylvester, and matrix square roots 50
4.4 Frobenius' proof of his square root theorem 52
5 The spread of Frobenius' ideas 53

We achieve this by writing $\|A\|_F = \sqrt{\langle A, A \rangle} = \sqrt{\mathrm{Tr}(A^T A)}$.

The Frobenius norm is the trace of a symmetric matrix

The 2×3 case is representative:

$$A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}, \|A\|_F = \sqrt{a^2 + b^2 + c^2 + d^2 + e^2 + f^2};$$

$$\begin{aligned} \mathrm{Tr}(A^T A) &= \mathrm{Tr}\left[\begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix} \times \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \right] \\ &= \mathrm{Tr} \begin{pmatrix} a^2 + d^2 & ab + de & ac + df \\ ba + ed & b^2 + e^2 & bc + ef \\ ca + fd & cb + fe & c^2 + f^2 \end{pmatrix} \\ &= a^2 + d^2 + b^2 + e^2 + c^2 + f^2 \\ &= \|A\|_F^2. \end{aligned}$$

Singular values

Let A be a matrix of size $m \times n$ over the reals. $A^T A$ is an $m \times m$ symmetric matrix.

Take the eigenvalues of $A^T A$ in nonincreasing order. Suppose these are $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$.

Real symmetric matrices have real nonnegative eigenvalues, so we can specify the nonincreasing sequence of m nonnegative real values $\sigma_i = \sqrt{\lambda_i}, i = 1 \ldots m$. These are the **singular values** of A.

E.g. $A = \begin{pmatrix} 2 & 0 & 3 \\ 3 & 1 & -2 \end{pmatrix}$

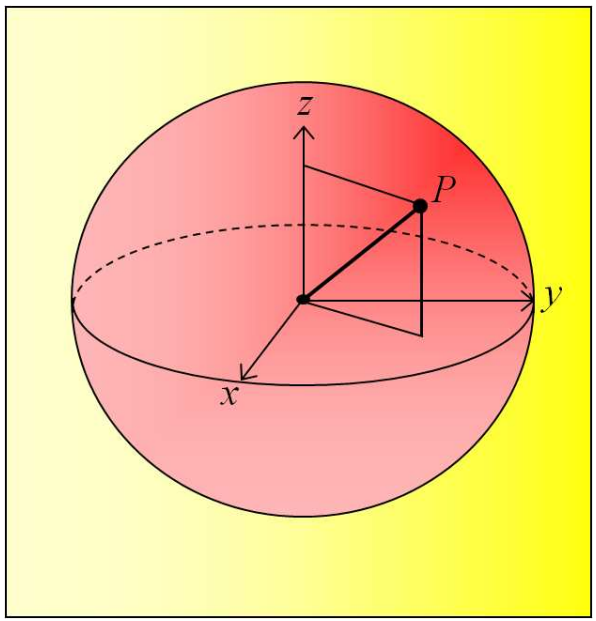
$$A^T A = \left[\begin{pmatrix} 2 & 3 \\ 0 & 1 \\ 3 & -2 \end{pmatrix} \times \begin{pmatrix} 2 & 0 & 3 \\ 3 & 1 & -2 \end{pmatrix} \right] = \begin{pmatrix} 13 & 3 & 0 \\ 3 & 1 & -2 \\ 0 & -2 & 13 \end{pmatrix}.$$

The eigenvalues of $A^T A$ are $14, 13, 0$; the singular values are $\sqrt{14}$, $\sqrt{13}$, 0.

Meaning of singular values

Let A be a matrix of size $m \times n$ over the reals, with singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_m$. Then the maximum value of $\|Ax\|$, where x ranges over unit vectors in $\mathbb{R}^n$, is the largest singular value σ_1, and this is achieved when x is an eigenvector of $A^T A$ with eigenvalue σ_1^2.

Think of our sphere as being radius 1 with centre at the origin. Point P is one of the unit vectors.



Suppose A is 3×2. As a linear operator this will create a stretching of a projection of the unit sphere in two dimensions. We want to know how much it is stretched and in which directions.

Singular values example

E.g. For a point $x = (x_1, x_2, x_3)$ on the unit sphere, what is the largest value of

$$\|Ax\| = \left\| \begin{pmatrix} 2 & 0 & 3 \\ 3 & 1 & -2 \end{pmatrix} \times (x_1, x_2, x_3)^T \right\| ?$$

The largest singular value is $\sqrt{14} \approx 3.74$. The unit vector which is stretched this far is the unit eigenvector corresponding to eigenvalue 14 of $A^T A$.

$$\begin{pmatrix} 13 & 3 & 0 \\ 3 & 1 & -2 \\ 0 & -2 & 13 \end{pmatrix} \times \frac{1}{\sqrt{14}} \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix} = \frac{1}{\sqrt{14}} \begin{pmatrix} 42 \\ 14 \\ -28 \end{pmatrix}.$$

Indeed,

$$\begin{pmatrix} 2 & 0 & 3 \\ 3 & 1 & -2 \end{pmatrix} \times \frac{1}{\sqrt{14}} \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix} = \frac{1}{\sqrt{14}} \begin{pmatrix} 0 \\ 14 \end{pmatrix}.$$

Secondary singular values

The second largest singular value, σ_2, is the maximum value of $\|Ax\|$ over unit vectors *orthogonal to the eigenvector of* σ_1.

E.g. Continuing our analysis of

$$\|Ax\| = \left\| \begin{pmatrix} 2 & 0 & 3 \\ 3 & 1 & -2 \end{pmatrix} \times (x_1, x_2, x_3)^T \right\|,$$

the second largest singular value of A is $\sqrt{13}$. The unit vector which is stretched this far, othogonal to $(0,1)^T$, is the unit eigenvector corresponding to eigenvalue 13 of A^TA.

$$\begin{pmatrix} 13 & 3 & 0 \\ 3 & 1 & -2 \\ 0 & -2 & 13 \end{pmatrix} \times \frac{2}{\sqrt{13}} \begin{pmatrix} 1 \\ 0 \\ 3/2 \end{pmatrix} = \frac{2}{\sqrt{13}} \begin{pmatrix} 13 \\ 0 \\ 39/2 \end{pmatrix}.$$

Indeed,

$$\begin{pmatrix} 2 & 0 & 3 \\ 3 & 1 & -2 \end{pmatrix} \times \frac{2}{\sqrt{13}} \begin{pmatrix} 1 \\ 0 \\ 3/2 \end{pmatrix} = \frac{2}{\sqrt{13}} \begin{pmatrix} 13/2 \\ 0 \end{pmatrix}.$$

Singular value decomposition

Let A be an $m \times n$ matrix of rank r. Then A can be written as a product

$$U\Sigma V^T,$$

where

- U is an $m \times m$ orthogonal matrix.
- V is an $n \times n$ orthogonal matrix.
- Σ is an $m \times n$ matrix whose ith diagonal entry equals the ith singular value σ_i for $i = 1, \ldots, r$. All other entries of Σ are zero.

The columns of V are orthonormal eigenvectors $v_1, \ldots, v_n$ of $A^T A$, where $A^T A v_i = \sigma_i^2 v_i$.
If $i \leq r$, so that $\sigma_i \neq 0$, then the ith column of U is $\sigma_i^{-1} A v_i$. These columns are orthonormal, and the remaining columns of U are obtained by arbitrarily extending to an orthonormal basis for $\mathbb{R}^m$.

SVD example

Continuing with our example $A = \begin{pmatrix} 2 & 0 & 3 \\ 3 & 1 & -2 \end{pmatrix}$.

We know that $\Sigma = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{pmatrix} = \begin{pmatrix} \sqrt{14} & 0 & 0 \\ 0 & \sqrt{13} & 0 \end{pmatrix}$.

Columns of V corresponding to σ_1 and σ_2 are, respectively,

$\frac{1}{\sqrt{14}} \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix}, \frac{2}{\sqrt{13}} \begin{pmatrix} 1 \\ 0 \\ 3/2 \end{pmatrix}$, $\sigma_3 = 0$ appends $\sqrt{\frac{2}{91}} \begin{pmatrix} -3/2 \\ 13/2 \\ 1 \end{pmatrix}$.

And already $\Sigma V^T = \begin{pmatrix} 3 & 1 & -2 \\ 2 & 0 & 3 \end{pmatrix}$.

So premultiplying by U just has to permute rows. And indeed, the columns of U are:

$\frac{1}{\sqrt{14}} Av_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\frac{1}{\sqrt{13}} Av_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

SVD calculator screen 1/3

You can experiment online with a (probably AI-generated) calculator (www.emathhelp.net/calculators/linear-algebra/svd-calculator/):

YOUR INPUT

Find the SVD of $\begin{bmatrix} 2 & 0 & 3 \\ 3 & 1 & -2 \end{bmatrix}$.

SOLUTION

Find the transpose of the matrix: $\begin{bmatrix} 2 & 0 & 3 \\ 3 & 1 & -2 \end{bmatrix}^T = \begin{bmatrix} 2 & 3 \\ 0 & 1 \\ 3 & -2 \end{bmatrix}$ (for steps, see **matrix transpose calculator**).

Multiply the transposed matrix with the initial: $Y = \begin{bmatrix} 2 & 3 \\ 0 & 1 \\ 3 & -2 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 & 3 \\ 3 & 1 & -2 \end{bmatrix} = \begin{bmatrix} 13 & 3 & 0 \\ 3 & 1 & -2 \\ 0 & -2 & 13 \end{bmatrix}$ (for steps, see **matrix multiplication calculator**).

Now, find the eigenvalues and eigenvectors of Y (for steps, see **eigenvalues and eigenvectors calculator**).

Eigenvalue: 14, **eigenvector:** $\begin{bmatrix} -\frac{3}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}$.

Eigenvalue: 13, **eigenvector:** $\begin{bmatrix} \frac{2}{3} \\ 0 \\ 1 \end{bmatrix}$.

(www.emathhelp.net/calculators/linear-algebra/svd-calculator/)

Note that these are scalar multiples of my choice of eigenvectors.

SVD calculator screen 2/3

... continuing with the third of the three eigenvectors:

Eigenvalue: 0, **eigenvector:** $\begin{bmatrix} -\frac{3}{2} \\ \frac{13}{2} \\ 1 \end{bmatrix}$.

Find the square roots of the nonzero eigenvalues (σ_i):

$\sigma_1 = \sqrt{14}$

$\sigma_2 = \sqrt{13}$

The Σ matrix is a zero matrix with σ_i on its diagonal: $\Sigma = \begin{bmatrix} \sqrt{14} & 0 & 0 \\ 0 & \sqrt{13} & 0 \end{bmatrix}$.

The columns of the matrix V are the normalized (unit) vectors: $V = \begin{bmatrix} -\frac{3\sqrt{14}}{14} & \frac{2\sqrt{13}}{13} & -\frac{3\sqrt{182}}{182} \\ -\frac{\sqrt{14}}{14} & 0 & \frac{\sqrt{182}}{14} \\ \frac{\sqrt{14}}{7} & \frac{3\sqrt{13}}{13} & \frac{\sqrt{182}}{91} \end{bmatrix}$ (for steps in finding a unit vector, see **unit vector calculator**).

Now, $u_i = \frac{1}{\sigma_i} \cdot \begin{bmatrix} 2 & 0 & 3 \\ 3 & 1 & -2 \end{bmatrix} \cdot v_i$:

$u_1 = \frac{1}{\sigma_1} \cdot \begin{bmatrix} 2 & 0 & 3 \\ 3 & 1 & -2 \end{bmatrix} \cdot v_1 = \frac{1}{\sqrt{14}} \cdot \begin{bmatrix} 2 & 0 & 3 \\ 3 & 1 & -2 \end{bmatrix} \cdot \begin{bmatrix} -\frac{3\sqrt{14}}{14} \\ -\frac{\sqrt{14}}{14} \\ \frac{\sqrt{14}}{7} \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ (for steps, see **matrix scalar multiplication calculator** and **matrix multiplication calculator**).

The columns of U change because of the different eigenvectors.

SVD calculator screen 3/3

... concluding with the $U\Sigma V^T$ decomposition:

$u_2 = \frac{1}{\sigma_2} \cdot \begin{bmatrix} 2 & 0 & 3 \\ 3 & 1 & -2 \end{bmatrix} \cdot v_2 = \frac{1}{\sqrt{13}} \cdot \begin{bmatrix} 2 & 0 & 3 \\ 3 & 1 & -2 \end{bmatrix} \cdot \begin{bmatrix} \frac{2\sqrt{13}}{13} \\ 0 \\ \frac{3\sqrt{13}}{13} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (for steps, see matrix scalar multiplication calculator and matrix multiplication calculator).

Therefore, $U = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

The matrices U, Σ, and V are such that the initial matrix $\begin{bmatrix} 2 & 0 & 3 \\ 3 & 1 & -2 \end{bmatrix} = U\Sigma V^T$.

ANSWER

$$U = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sqrt{14} & 0 & 0 \\ 0 & \sqrt{13} & 0 \end{bmatrix} \approx \begin{bmatrix} 3.741657386773941 & 0 & 0 \\ 0 & 3.605551275463989 & 0 \end{bmatrix}$$

$$V = \begin{bmatrix} -\frac{3\sqrt{14}}{14} & \frac{2\sqrt{13}}{13} & -\frac{3\sqrt{182}}{182} \\ -\frac{\sqrt{14}}{14} & 0 & \frac{\sqrt{182}}{14} \\ \frac{\sqrt{14}}{7} & \frac{3\sqrt{13}}{13} & \frac{\sqrt{182}}{91} \end{bmatrix} \approx \begin{bmatrix} -0.801783725737273 & 0.554700196225229 & -0.22237479499833 \\ -0.267261241912424 & 0 & 0.963624111659432 \\ 0.534522483824849 & 0.832050294337844 & 0.14824986333222 \end{bmatrix}$$

So SVD is not unique. The singular values are uniquely determined.

Singular values and the Frobenius norm

We write $A = U\Sigma V^T$, U, V orthogonal (transpose is inverse). Multiplying vectors by an orthogonal matrix does not change the Euclidean norm of the vector (orthogonal matrices form a group which is precisely the group of distance-preserving transformations). So we have

$$\|A\|_F^2 = \|\Sigma\|_F^2 = \sigma_1^2 + \sigma_2^2 + \ldots + \sigma_m^2.$$

Here is a quote from a (probably AI-generated) explanation:

> *This formula illuminates what the Frobenius norm is really measuring: it is the total energy spread across all the "modes" of the matrix, where each singular value σ_i is the amplitude of the ith mode. A matrix concentrated in a single mode has most of its Frobenius norm in σ_1^2; a matrix that spreads energy evenly across many directions will have many comparable singular values.*
>
> lessonsvault.app/lessons/mathematics/frobenius-norm/

SVD and other norms

There appear to be at least two other matrix norms on offer for a matrix A :

Spectral norm: $\|A\|_2 = \sigma_1$, the largest singular value of A.

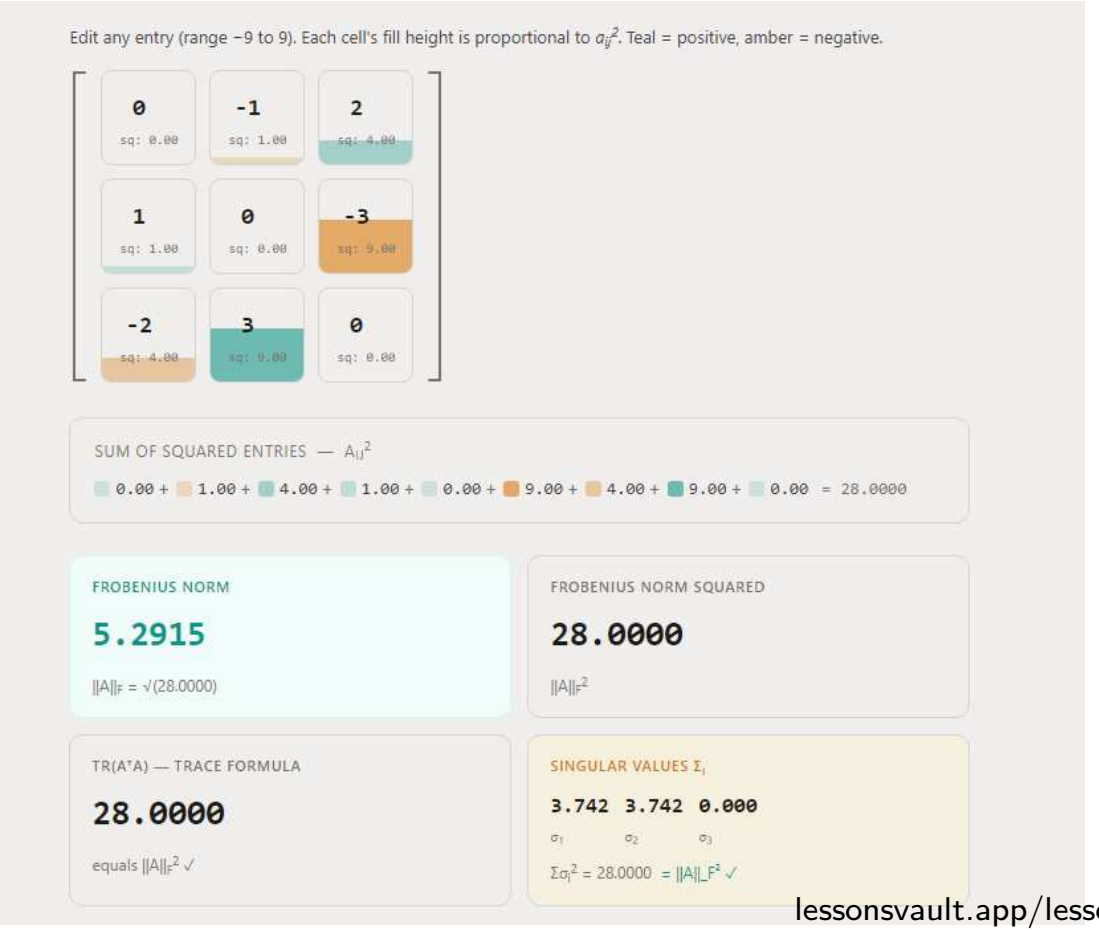
Nuclear norm: $\|A\|_* = \sum \sigma_i$, the sum of the singular values of A.

I don't know anything about these but apparently they satisfy some pretty inequalities:

$$\begin{aligned} \|A\|_2 &\leq \|A\|_F \leq \sqrt{r}\,\|A\|_2, \\ \|A\|_F &\leq \|A\|_* \leq \sqrt{r}\,\|A\|_F, \end{aligned}$$

where r is the rank of A.

The Frobenius norm gets an online calculator too!



lessonsvault.app/lessons/mathematics/frobenius-norm/

This is the image on my title slide. The calculator only seems to do 3×3 matrices but this is OK because my interest in the Frobenius norm doesn't go beyond square matrices. In fact, my interest is specifically in skew-symmetric matrices ($A^T = -A$), like the one shown.

Frobenius norms for skew-symmetric matrices

For a matrix X denote by $S(X)$ the sum of the elements of X

Let A be an $n \times n$ skew-symmetric matrix.

Let $\hat{A}$ be the $(n+1) \times (n+1)$ matrix formed from A by adding an $n+1$th row and column having minus the column (respectively row) sums of A.

E.g. $A = \begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & -3 \\ -2 & 3 & 0 \end{pmatrix}$, $\hat{A} = \begin{pmatrix} 0 & -1 & 2 & -1 \\ 1 & 0 & -3 & 2 \\ -2 & 3 & 0 & -1 \\ 1 & -2 & 1 & 0 \end{pmatrix}$

Then

$$\|\hat{A}\|_F = \|A\|_F + 2S(A^2).$$

This is maybe only interesting (only to me) because it's a step on the way to linking my triangle area matrices.

As well as the (probably AI-generated) webpages cited in previous slides I found some lecture notes by Michael Hutchings at Berkeley very clear and helpful: math.berkeley.edu/∼hutching/teach/54-2017/svd-notes.pdf