The Singular Value Decomposition for Analysts

Jim Rulla

January 13, 2023

For $\mathsf{x} \in R^n$ and $\mathsf{y} \in R^m$,

$$\|\mathsf{x}\| = \sqrt{\sum_{j=1}^{n} x_j^2} \qquad \|\mathsf{y}\| = \sqrt{\sum_{i=1}^{m} y_i^2}$$

$\mathsf{A} : R^n \to R^m$ is linear

$$\|\mathsf{A}\| = \sup_{\|\mathsf{x}\| \leq 1} \|\mathsf{A}\mathsf{x}\| = \sup_{\|\mathsf{x}\| = 1} \|\mathsf{A}\mathsf{x}\|$$

Example: Consider

$$\mathsf{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 - x_1 \\ x_2 \\ x_1 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

$$\begin{aligned} \left\| \mathsf{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\|^2 &= (x_2 - x_1)^2 + x_2^2 + x_1^2 \\ &= 2\left(x_1^2 - x_1 x_2 + x_2^2\right). \end{aligned}$$

Parameterize the unit vector

$$\mathsf{x}(\theta) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix},$$

so

$$\|\mathsf{A}\mathsf{x}(\theta)\|^2 = 2 - \sin(2\theta).$$

The most A can stretch a unit vector is

$$\|\mathsf{A}\| = \sup_{\theta \in [0, 2\pi)} \sqrt{2 - \sin(2\theta)} = \sqrt{3}.$$

$\theta = \frac{3\pi}{4}$ maximizes $\|\mathsf{A}\mathsf{x}(\theta)\| = \sqrt{2 - \sin(2\theta)}$, so a unit vector which A stretches the most is

$$\mathsf{v}_1 = \begin{bmatrix} \cos\left(\frac{3\pi}{4}\right) \\ \sin\left(\frac{3\pi}{4}\right) \end{bmatrix} = \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix},$$

and its corresponding image $\mathsf{A}\mathsf{v}_1$ is

$$\begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \mathsf{v}_1 = \begin{bmatrix} \sqrt{2} \\ \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{bmatrix} = \underbrace{\sqrt{3}}_{\sigma_1} \underbrace{\begin{bmatrix} 2\frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} \\ -\frac{\sqrt{6}}{6} \end{bmatrix}}_{\mathsf{u}_1}.$$

v_1 is a **right singular vector**,
u_1 is the corresponding **left singular vector**,
$\sigma_1 = \|\mathsf{A}\| = \sqrt{3}$ is A's **largest singular value**.

Wait, there's more!

$\theta = \frac{\pi}{4}$ *minimizes* $\|\mathsf{A}\mathsf{x}(\theta)\| = \sqrt{2 - \sin(2\theta)}$, so a unit vector which A stretches the *least* is

$$\mathsf{v}_2 = \begin{bmatrix} \cos\left(\frac{\pi}{4}\right) \\ \sin\left(\frac{\pi}{4}\right) \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix},$$

and the corresponding image $\mathsf{A}\mathsf{v}_2$ is

$$\begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \mathsf{v}_2 = \begin{bmatrix} 0 \\ \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} = \underbrace{1}_{\sigma_2} \underbrace{\begin{bmatrix} 0 \\ \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}}_{\mathsf{u}_2}.$$

Observe that $\mathsf{v}_1 \perp \mathsf{v}_2$ (from $\sin(2\theta)$, *e.g.*) and $\mathsf{u}_1 \perp \mathsf{u}_2$ (from dumb luck, or not?).

v_1 & v_2 are an orthonormal basis of R^2, so $\forall\, x$,

$$x = v_1 {v_1}^T x + v_2 {v_2}^T x,$$
$$Ax = \sigma_1 u_1 {v_1}^T x + \sigma_2 u_2 {v_2}^T x,$$

$$\text{So } A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T$$
$$= \begin{bmatrix} | & | \\ u_1 & u_2 \\ | & | \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} - & {v_1}^T & - \\ - & {v_2}^T & - \end{bmatrix}$$
$$+ \begin{bmatrix} | & | \\ u_1 & u_2 \\ | & | \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} - & {v_1}^T & - \\ - & {v_2}^T & - \end{bmatrix}$$
$$= \underbrace{\begin{bmatrix} | & | \\ u_1 & u_2 \\ | & | \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} - & {v_1}^T & - \\ - & {v_2}^T & - \end{bmatrix}}_{V^T}$$

is a **Singular Value Decomposition** of A.

In our example,

$$\begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2\frac{\sqrt{6}}{6} & 0 \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

U has orthonormal columns, so $U^T U = I_{2\times 2}$ and UU^T = orthogonal projection onto $\langle u_1, u_2 \rangle$.

V has orthonormal columns, so $V^T V = I_{2\times 2}$ and VV^T = orthogonal projection onto $\langle v_1, v_2 \rangle$.

Σ is diagonal, with non-negative entries.

yx^T is an **outer product**, a **rank-1 matrix** or a **dyad**.

The SVD represents A as a linear combination of two rank-1 matrices, both of whose "y"s and both of whose "x"s are orthonormal.

Lemma: $\mathcal{V}, \mathcal{U}$ Hilbert spaces, $\mathsf{A} : \mathcal{V} \to \mathcal{U}$ linear. If the unit vector v_0 satisfies

$$\|\mathsf{A}\| = \|\mathsf{A}\mathsf{v}_0\|,$$

then

$$\mathsf{A}\mathsf{v}_1 \perp \mathsf{A}\mathsf{v}_0 \quad \text{whenever} \quad \mathsf{v}_1 \perp \mathsf{v}_0.$$

In other words, $\mathsf{A}\langle \mathsf{v}_0 \rangle^\perp \leq \langle \mathsf{A}\mathsf{v}_0 \rangle^\perp$.

Lemma's lemma ("Complete the cosine"):

$$a\cos(\alpha) + b\sin(\alpha) = r\cos(\alpha - \varphi),$$

where

$$r = \sqrt{a^2 + b^2}, \quad \cos(\varphi) = \frac{a}{r}, \quad \& \quad \sin(\varphi) = \frac{b}{r}.$$

Proof of the lemma: v_1 may as well be a unit vector. Since $\mathsf{v}_1 \perp \mathsf{v}_0$,

$$\mathsf{v}(\theta) = \cos(\theta)\mathsf{v}_0 + \sin(\theta)\mathsf{v}_1$$

is a unit vector for every θ.

$$\begin{aligned}
\|\mathsf{A}\mathsf{v}(\theta)\|^2 &= \cos^2(\theta)\,\|\mathsf{A}\mathsf{v}_0\|^2 \\
&\quad + 2\sin(\theta)\cos(\theta)\,(\mathsf{A}\mathsf{v}_0, \mathsf{A}\mathsf{v}_1) \\
&\quad + \sin^2(\theta)\,\|\mathsf{A}\mathsf{v}_1\|^2 \\
&= \tfrac{\|\mathsf{A}\mathsf{v}_0\|^2 + \|\mathsf{A}\mathsf{v}_1\|^2}{2} + \tfrac{\|\mathsf{A}\mathsf{v}_0\|^2 - \|\mathsf{A}\mathsf{v}_1\|^2}{2}\cos(2\theta) \\
&\quad + (\mathsf{A}\mathsf{v}_0, \mathsf{A}\mathsf{v}_1)\sin(2\theta) \\
&= \tfrac{\|\mathsf{A}\mathsf{v}_0\|^2 + \|\mathsf{A}\mathsf{v}_1\|^2}{2} + r\cos(2\theta - \varphi),
\end{aligned}$$

where, by the lemma's lemma,

$$r = \sqrt{\left(\tfrac{\|\mathsf{A}\mathsf{v}_0\|^2 - \|\mathsf{A}\mathsf{v}_1\|^2}{2}\right)^2 + (\mathsf{A}\mathsf{v}_0, \mathsf{A}\mathsf{v}_1)^2}.$$

So

$$\|\mathsf{Av}_0\|^2 = \frac{\|\mathsf{Av}_0\|^2 + \|\mathsf{Av}_1\|^2}{2} + \sqrt{\left(\frac{\|\mathsf{Av}_0\|^2 - \|\mathsf{Av}_1\|^2}{2}\right)^2 + (\mathsf{Av}_0, \mathsf{Av}_1)^2} > \|\mathsf{Av}_0\|^2 \quad \text{if} \quad (\mathsf{Av}_0, \mathsf{Av}_1) \neq 0.$$

////

The proof holds for any Hilbert spaces, not just finite-dimensional spaces.

The lemma is a differential condition for extrema of a function (the norm) subject to a constraint (to unit vectors): at a maximizer v_0, movement in any allowed direction (perpendicular to v_0) must not increase the length of the image u_0; hence must move orthogonal to the image.

If $\|\mathsf{v}(\theta)\|^2 = 1$, then

$$0 = \frac{d}{d\theta} \|\mathsf{v}(\theta)\|^2 = 2\mathsf{v}(\theta)^T \mathsf{v}'(\theta),$$

so the derivative is orthogonal to the vector.

Likewise, at an extremum,

$$0 = \frac{d}{d\theta} \|\mathsf{Av}(\theta)\|^2 = 2\mathsf{v}(\theta)^T \mathsf{A}^T \mathsf{Av}'(\theta),$$

so Av_0 is orthogonal to the image of $\mathsf{v}'(\theta)$.

We construct the SVD of A recursively.

Step 1: Find a unit vector v_1 for which

$$\|\mathsf{A}\| = \|\mathsf{A}\mathsf{v}_1\| = \sigma_1,$$

and a normalized image u_1 for which

$$\mathsf{A}\mathsf{v}_1 = \sigma_1 \mathsf{u}_1.$$

By the lemma, A maps

$$\begin{array}{rcl} \mathcal{S}_1 = \langle \mathsf{v}_1 \rangle & \longrightarrow & \langle \mathsf{u}_1 \rangle \\ \mathcal{S}_1{}^{\perp} = \langle \mathsf{v}_1 \rangle^{\perp} & \longrightarrow & \langle \mathsf{u}_1 \rangle^{\perp} \end{array}$$

Step 2: $\mathcal{S}_1{}^{\perp} = \langle \mathsf{v}_1 \rangle^{\perp}$ is a Hilbert space. Find a unit vector $\mathsf{v}_2 \in \mathcal{S}_1{}^{\perp}$ for which

$$\sup_{\substack{\mathsf{v} \in \mathcal{S}_1{}^{\perp} \\ \|\mathsf{v}\|=1}} \|\mathsf{A}\mathsf{v}\| = \|\mathsf{A}\mathsf{v}_2\| = \sigma_2,$$

and a normalized image u_2 for which

$$\mathsf{A}\mathsf{v}_2 = \sigma_2 \mathsf{u}_2.$$

Since $\langle \mathsf{v}_2 \rangle^{\perp}$ in $\mathcal{S}_1{}^{\perp}$ is $\langle \mathsf{v}_1, \mathsf{v}_2 \rangle^{\perp}$, the lemma says A maps

$$\begin{array}{rcl} \mathcal{S}_2 = \langle \mathsf{v}_1, \mathsf{v}_2 \rangle & \longrightarrow & \langle \mathsf{u}_1, \mathsf{u}_2 \rangle \\ \mathcal{S}_2{}^{\perp} = \langle \mathsf{v}_1, \mathsf{v}_2 \rangle^{\perp} & \longrightarrow & \langle \mathsf{u}_1, \mathsf{u}_2 \rangle^{\perp} \end{array}$$

Furthermore, A stretches every vector in $\langle \mathsf{v}_1, \mathsf{v}_2 \rangle$ by at least σ_2:

$$\begin{aligned} \|\mathsf{A}(\alpha_1 \mathsf{v}_1 + \alpha_2 \mathsf{v}_2)\|^2 &= \|\alpha_1 \sigma_1 \mathsf{u}_1 + \alpha_2 \sigma_2 \mathsf{u}_2\|^2 \\ &= \sigma_1^2 \|\alpha_1 \mathsf{u}_1\|^2 + \sigma_2^2 \|\alpha_2 \mathsf{u}_2\|^2 \\ &\geq \sigma_2^2 \left(\alpha_1^2 + \alpha_2^2\right) \\ &= \sigma_2^2 \|\alpha_1 \mathsf{v}_1 + \alpha_2 \mathsf{v}_2\|^2 \end{aligned}$$

because $\mathsf{u}_1 \perp \mathsf{u}_2$, $\sigma_1 \geq \sigma_2$, and $\mathsf{v}_1 \perp \mathsf{v}_2$.

Step k: Set $\mathcal{S}_{k-1} = \langle \mathsf{v}_1, \cdots, \mathsf{v}_{k-1} \rangle$. Then ${\mathcal{S}_{k-1}}^\perp$ is a Hilbert space. Find a unit vector $\mathsf{v}_k \in {\mathcal{S}_{k-1}}^\perp$ for which

$$\sup_{\substack{\mathsf{v} \in {\mathcal{S}_{k-1}}^\perp \\ \|\mathsf{v}\|=1}} \|\mathsf{A}\mathsf{v}\| = \|\mathsf{A}\mathsf{v}_k\| = \sigma_k,$$

and a normalized image u_k for which

$$\mathsf{A}\mathsf{v}_k = \sigma_k \mathsf{u}_k.$$

By the lemma, A maps

$$\begin{array}{rcl} \mathcal{S}_k = \langle \mathsf{v}_1, \ldots, \mathsf{v}_k \rangle & \longrightarrow & \langle \mathsf{u}_1, \ldots, \mathsf{u}_k \rangle \\ {\mathcal{S}_k}^\perp = \langle \mathsf{v}_1, \ldots, \mathsf{v}_k \rangle^\perp & \longrightarrow & \langle \mathsf{u}_1, \ldots, \mathsf{u}_k \rangle^\perp \end{array}$$

Furthermore, $\sigma_k \le \sigma_{k-1}$,

$$\begin{aligned} \|\mathsf{A}\mathsf{v}\| &\ge \sigma_k \|\mathsf{v}\| \quad \forall\, \mathsf{v} \in \langle \mathsf{v}_1, \ldots, \mathsf{v}_k \rangle, \text{ and} \\ \|\mathsf{A}\mathsf{v}\| &\le \sigma_k \|\mathsf{v}\| \quad \forall\, \mathsf{v} \in \langle \mathsf{v}_1, \ldots, \mathsf{v}_k \rangle^\perp. \end{aligned}$$

Iterate until $\sigma_{r+1} = 0$. Then

$$\begin{aligned} \|\mathsf{A}\mathsf{v}\| &\ge \sigma_r \|\mathsf{v}\| \quad \forall\, \mathsf{v} \in \mathcal{S}_r = \langle \mathsf{v}_1, \ldots, \mathsf{v}_r \rangle \\ \|\mathsf{A}\mathsf{v}\| &= 0 \quad \forall\, \mathsf{v} \in {\mathcal{S}_r}^\perp = \langle \mathsf{v}_1, \ldots, \mathsf{v}_r \rangle^\perp \end{aligned}$$

so

$$\langle \mathsf{v}_1, \ldots, \mathsf{v}_r \rangle^\perp = \mathcal{N}(\mathsf{A}).$$

For any x,

$$\mathsf{x} = \mathsf{v}_1 \mathsf{v}_1^T \mathsf{x} + \cdots + \mathsf{v}_r \mathsf{v}_r^T \mathsf{x} + \mathsf{n},$$

where $\mathsf{n} \in \mathcal{N}(\mathsf{A})$. Then

$$\mathsf{A}\mathsf{x} = \mathsf{u}_1 \sigma_1 \mathsf{v}_1^T \mathsf{x} + \cdots + \mathsf{u}_r \sigma_r \mathsf{v}_r^T \mathsf{x}, = \mathsf{U}\Sigma\mathsf{V}^T \mathsf{x}$$

so an SVD exists,

$\langle \mathsf{u}_1, \ldots, \mathsf{u}_r \rangle$ is A's **range** or **column space**, and r is the **rank** of A.

SVDs represent operators as sums of rank-1 operators:

$$\mathsf{A} = \sum_{i=1}^{r} \mathsf{u}_i \sigma_i \mathsf{v}_i^T.$$

The "bases" $\{\mathsf{v}_i\}$ and $\{\mathsf{u}_i\}$ are orthonormal, so term i is a (scalar) projection onto v_i, "stretched" by σ_i, "pointing" in the direction of u_i.

Since the σ_i are non-increasing,

$$\|\mathsf{A}\mathsf{v}\| \geq \sigma_k \|\mathsf{v}\| \; \forall \, \mathsf{v} \in \langle \mathsf{v}_1, \ldots, \mathsf{v}_k \rangle$$
$$\|\mathsf{A}\mathsf{v}\| \leq \sigma_k \|\mathsf{v}\| \; \forall \, \mathsf{v} \in \langle \mathsf{v}_1, \ldots, \mathsf{v}_k \rangle^{\perp}$$

The most "rank-1" information about A is contained in $\mathsf{u}_1\sigma_1\mathsf{v}_1^T$. The most "rank-2" information is contained in $\mathsf{u}_1\sigma_1\mathsf{v}_1^T + \mathsf{u}_2\sigma_2\mathsf{v}_2^T$, etc.

Schmidt sought low-rank approximations of operators.

The SVD has a "tidy representation" as

$$\mathsf{A} = \mathsf{U}\Sigma\mathsf{V}^T = \begin{bmatrix} | & & | \\ \mathsf{u}_1 & \cdots & \mathsf{u}_r \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{bmatrix} \begin{bmatrix} -{\mathsf{v}_1}^T- \\ \vdots \\ -{\mathsf{v}_r}^T- \end{bmatrix}$$

where $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$.

Since the v_i are orthonormal, $\mathsf{V}\mathsf{V}^T$ is the orthogonal projection onto the span of the v_i, and similarly for $\mathsf{U}\mathsf{U}^T$. Also, $\mathsf{V}^T\mathsf{V} = \mathsf{I}_{r\times r} = \mathsf{U}^T\mathsf{U}$.

The SVD is especially useful in applications requiring the rank of A.

If you already know the spectral theorem, you can use Lagrange Multipliers to find v_i:

$$\nabla \|\mathsf{A}\mathsf{v}_i\|^2 = \lambda_i \nabla \|\mathsf{v}_i\|^2 \Rightarrow \mathsf{A}^T\mathsf{A}\mathsf{v}_i = \lambda_i \mathsf{v}_i.$$

Left multiply by v_i^T to verify $\lambda_i = \|\mathsf{A}\mathsf{v}_i\|^2 = \sigma_i^2$.

Since

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

an SVD for the transpose is

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

$\langle \mathsf{u}_1, \ldots, \mathsf{u}_r \rangle^\perp$ is the **left null space** of A, and

$\langle \mathsf{v}_1, \ldots, \mathsf{v}_r \rangle$ is the **row space**.

The rank of A^T is r, the rank of A.

These conclusions are collectively known as the **Fundamental Theorem of Linear Algebra**.

Furthermore, the singular values of A^T are the same as the singular values of A.

The maximum- and minimum-stretching properties of the singular values imply the "δ-ϵ" estimates

$$\|\mathsf{A}\mathsf{x}_2 - \mathsf{A}\mathsf{x}_1\| \leq \sigma_1 \|\mathsf{x}_2 - \mathsf{x}_1\|$$

for all x_1 and x_2, and

$$\sigma_r \|\mathsf{x}_2 - \mathsf{x}_1\| \leq \|\mathsf{A}\mathsf{x}_2 - \mathsf{A}\mathsf{x}_1\|$$

for all x_1 and x_2 in the row space of A.

A is one-to-one on the row space, so it is one-to-one if $r = n$.

If A is one-to-one, then it maps any region of (n-dim'l) volume V to a region of (n-dim'l) volume $\prod \sigma_i V$.

If A is onto, then it maps a δ-ball into a $\sigma_r\delta$ ball. This is a "hard" Open Mapping Theorem.

Suppose we know x to one part in a thousand. How much relative accuracy can we ascribe to Ax? In other words, if

$$\frac{\|\mathsf{h}\|}{\|\mathsf{x}\|} < \frac{1}{1000},$$

what can we say about

$$\frac{\|\mathsf{A}(\mathsf{x}+\mathsf{h}) - \mathsf{Ax}\|}{\|\mathsf{Ax}\|}\,?$$

If $\mathsf{Ax} = \mathsf{0}$, then the notion of relative accuracy is moot. Suppose, therefore, that $r = n$, so that $\mathcal{N}(\mathsf{A}) = \{\mathsf{0}\}$. Then the singular values of A give the solution straight away:

$$\frac{\|\mathsf{A}(\mathsf{x}+\mathsf{h}) - \mathsf{Ax}\|}{\|\mathsf{Ax}\|} = \frac{\|\mathsf{Ah}\|}{\|\mathsf{Ax}\|} \le \frac{\sigma_1}{\sigma_r}\frac{\|\mathsf{h}\|}{\|\mathsf{x}\|} \le \frac{\sigma_1}{\sigma_r}\frac{1}{1000}$$

The ratio

$$\kappa = \frac{\sigma_1}{\sigma_r}$$

is the **condition number** of A. If we know x to one part in a 1000, then we know Ax to κ parts in a 1000.

The smallest condition number is $\kappa = 1$. Operators with κ near 1 are “well-conditioned”. They map the unit sphere to a very-nearly-spherical ellipsoid.

Operators with large κ are “ill-conditioned”. They stretch one direction much more than another.

Ill-conditioning doesn't matter much in the direction of the "largest" singular vector v_1:

$$\frac{\|\mathsf{A}(\mathsf{v}_1 + \mathsf{h}) - \mathsf{A}\mathsf{v}_1\|}{\|\mathsf{A}\mathsf{v}_1\|} = \frac{\|\mathsf{A}\mathsf{h}\|}{\|\sigma_1 \mathsf{u}_1\|} \leq \frac{\sigma_1 \|\mathsf{h}\|}{\sigma_1} = \frac{\|\mathsf{h}\|}{\|\mathsf{v}_1\|}.$$

In the direction of the "smallest" singular vector v_r, however, ill-conditioning can be disastrous:

$$\frac{\|\mathsf{A}(\mathsf{v}_r + \mathsf{h}) - \mathsf{A}\mathsf{v}_r\|}{\|\mathsf{A}\mathsf{v}_r\|} = \frac{\|\mathsf{A}\mathsf{h}\|}{\|\sigma_r \mathsf{u}_r\|} \leq \frac{\sigma_1 \|\mathsf{h}\|}{\sigma_r} = \kappa \frac{\|\mathsf{h}\|}{\|\mathsf{v}_r\|}.$$

Ill-conditioned operators magnify rounding errors in some directions much more than in others. If part of your answer is correct and part is not, check for ill-conditioning.

Gaussian elimination consists of operations of the form

$$\begin{bmatrix} a_{ii} & a_{i(i+1)} & \cdots \\ a_{ji} & a_{j(i+1)} & \cdots \end{bmatrix} \to \begin{bmatrix} a_{ii} & a_{i(i+1)} & \cdots \\ 0 & a_{j(i+1)} - \frac{a_{i(i+1)}a_{ji}}{a_{ii}} & \cdots \end{bmatrix}.$$

The operator for a step of elimination is

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

where $\alpha = -\frac{a_{ji}}{a_{ii}}$. As always, we compute

$$\left\| \mathsf{A} \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} \cos(\theta) \\ \alpha\cos(\theta) + \sin(\theta) \end{bmatrix} \right\|^2$$
$$= 1 + 2\alpha\cos(\theta)\sin(\theta) + \alpha^2\cos^2(\theta)$$
$$= 1 + \frac{\alpha^2}{2} + \alpha\sin(2\theta) + \frac{\alpha^2}{2}\cos(2\theta).$$

"Complete the cosine" to find

$$\sigma_1^2 = 1 + \frac{\alpha^2}{2} + \sqrt{\alpha^2 + \frac{\alpha^4}{4}}$$
$$\sigma_2^2 = 1 + \frac{\alpha^2}{2} - \sqrt{\alpha^2 + \frac{\alpha^4}{4}},$$

so the condition number is

$$\kappa = \sqrt{\frac{1 + \frac{\alpha^2}{2} + \sqrt{\alpha^2 + \frac{\alpha^4}{4}}}{1 + \frac{\alpha^2}{2} - \sqrt{\alpha^2 + \frac{\alpha^4}{4}}}}$$
$$= 1 + \frac{\alpha^2}{2} + \sqrt{\alpha^2 + \frac{\alpha^4}{4}}.$$

The trouble: $\kappa \sim \alpha^2$ if $\alpha = -\frac{a_{ji}}{a_{ii}}$ is large.

At the very least, numerical elimination code should arrange for $|\alpha| < 1$. That technique is called **pivoting**.

Suppose a rank-4 matrix has

$$\sigma_1 = 7, \quad \sigma_2 = 5, \quad \sigma_3 = 3, \quad \text{and} \quad \sigma_4 = 10^{-15}.$$

The condition number $\kappa = 7 \times 10^{15}$ is enormous. Numerically speaking, σ_4 probably has more to do with rounding errors than reality, and should probably have been zero. In this case, we say the **numerical rank** of A is 3.

Dropping the last term in

$$\mathsf{A} = 7\mathsf{u}_1\mathsf{v}_1^T + 5\mathsf{u}_2\mathsf{v}_2^T + 3\mathsf{u}_3\mathsf{v}_3^T + 10^{-15}\mathsf{u}_4\mathsf{v}_4^T$$

to get a rank-3 SVD is the same as setting $\sigma_4 = 0$, which is the same as stopping the construction of the SVD one step earlier.

Storage space for A requires mn values, while

$$\sum_{i=1}^{r} \sigma_i \mathsf{u}_i \mathsf{v}_i^T$$

requires $r(1 + m + n)$.

If the (numerical) rank r is much less than m and n, then the SVD requires much less storage than A. An array of pixels in a photograph might therefore be compressed using the SVD.

If $\mathsf{A} : R^n \to R^n$ is invertible, then U and V are orthogonal matrices, and an SVD of A^{-1} is

$$\mathsf{A}^{-1} = \left(\mathsf{U}\Sigma\mathsf{V}^T\right)^{-1} = \mathsf{V}\Sigma^{-1}\mathsf{U}^T. \qquad (PI)$$

Since Σ is diagonal with positive diagonal entries, its inverse is simple to compute.

Equation (PI) makes sense even if A is not invertible. The result is the **pseudoinverse**:

$$\mathsf{A}^{\dagger} = [\,\mathsf{v}_1\ \mathsf{v}_2\ \cdots\ \mathsf{v}_r\,] \begin{bmatrix} \frac{1}{\sigma_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_r} \end{bmatrix} \begin{bmatrix} \mathsf{u}_1^T \\ \mathsf{u}_2^T \\ \vdots \\ \mathsf{u}_r^T \end{bmatrix}.$$

$\mathsf{A}^{\dagger} : R^m \to R^n$ maps the span of the u_j, which is the range of A, to the span of the v_i, which is the row space in the domain of A. Furthermore,

$$\mathsf{A}^{\dagger}\mathsf{A} = \mathsf{V}\mathsf{V}^T \quad \text{and} \quad \mathsf{A}\mathsf{A}^{\dagger} = \mathsf{U}\mathsf{U}^T,$$

the orthogonal projections onto, respectively, the row space in the domain, and the range of A.

Remark: The condition number of $\mathsf{A}^{\dagger}$ is $\frac{\frac{1}{\sigma_r}}{\frac{1}{\sigma_1}} = \frac{\sigma_1}{\sigma_r} = \kappa$. If A is well-conditioned, then so is the pseudoinverse.
If A is ill-conditioned, then the best conditioned directions for A are the worst conditioned directions for the pseudoinverse.

The SVD is useful in least-squares problems.

Example: In the simplest case, we want to know whether data points $\begin{bmatrix} x_i \\ y_i \end{bmatrix}$ lie on a line $ax + by + c = 0$. The least-squares solution chooses a, b, and c to minimize

$$\sum_{i=1}^{m} (ax_i + by_i + c)^2 = \left\| \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots \\ x_m & y_m & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \right\|^2$$

over all possible a, b, and c. O.K., that's too easy: just take $a = b = c = 0$. I guess we really want to know if there is a *non-trivial* minimum. Suppose we require $a^2 + b^2 + c^2 = 1$. Then we want

$$\min_{a^2+b^2+c^2=1} \left\| \begin{bmatrix} x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots \\ x_m & y_m & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \right\|^2$$

which is just σ_3^2, square of the smallest singular value of the (data) matrix $\mathsf{X} = \begin{bmatrix} x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots \\ x_m & y_m & 1 \end{bmatrix}$.

The "smallest" right singular vector v_3 gives the optimal estimate for the line's coefficients,

$$\begin{bmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{bmatrix} = \mathsf{v}_3.$$

This is a problem in which the *smallest* singular value is what we seek first. The smaller σ_3 is, the better the line fits the data.

How do I go about designing a voice compression algorithm?

Fourier is frequently disappointing because it requires harmonics commensurate with frame length.

What if

$$s(t) = \sum_j a_j \cos(\omega_j t) + b_i \sin(\omega_j t)?$$

(By Slide 4's $a\cos(\alpha) + b\sin(\alpha) = r\cos(\alpha - \varphi)$, all phases of sinusoids of frequency ω_j are represented here!)

For digital compression, we sample a signal and compress the samples.

How many "harmonics" do we need?

The (numerical) rank of

$$\begin{bmatrix} s_0 & s_1 & s_2 & \cdots \\ s_1 & s_2 & s_3 & \cdots \\ s_2 & s_3 & s_4 & \cdots \\ s_3 & s_4 & s_5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

ought to be 2× # of "harmonics" present.

If the numerical rank is r, then any $r+1$ columns are linearly dependent:

$$c_0 s_k + c_1 s_{k+1} + \cdots + c_{r-1} s_{k+r-1} = s_{k+r}.$$

Find the c_j to predict the next sample from the previous r samples — Linear Predictive Coding.

LPC is a discrete linear ODE describing the signal. To compress sound, the encoder transmits the c_j and the decoder applies them to some initial values.