Plan for Today – and Tomorrow Morning

1. A bit of motivation.
2. Singular values and functions.
3. Picard condition and “range condition.”
4. The singular value decomposition (SVD) of a matrix.
5. SVD analysis.

Points to take home:
- Singular values/functions provide insight into the Radon transform:
  - High-frequency noise has a larger effect on the reconstruction.
  - Errors on the sinogram edges are especially troublesome.
- Filtering is necessary to reduce the influence of the noise.
- When the problem is discretized, we use the SVD to obtain insight.
- The SVD can always be used to study more general CT scenarios.
### Some Notation

<table>
<thead>
<tr>
<th>Vectors</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Norm</strong></td>
<td><strong>Functions</strong></td>
</tr>
<tr>
<td>$|x|<em>2^2 = \sum</em>{i=1}^{n}</td>
<td>x_i</td>
</tr>
<tr>
<td>$= x \cdot x$</td>
<td>$= \langle f, f \rangle$</td>
</tr>
<tr>
<td><strong>Inner prod.</strong></td>
<td><strong>Inner prod.</strong></td>
</tr>
<tr>
<td>$x \cdot y = \sum_{i=1}^{n} \bar{x}_i y_i = \bar{x}^T y$</td>
<td>$\langle f, g \rangle = \int_{a}^{b} \overline{f(t)} g(t) , dt$</td>
</tr>
<tr>
<td><strong>Orthonormal</strong></td>
<td><strong>Orthonormal</strong></td>
</tr>
<tr>
<td>$v_i \cdot v_j = \delta_{ij}$</td>
<td>$\langle v_i, v_j \rangle = \delta_{ij}$</td>
</tr>
<tr>
<td><strong>Expansion</strong></td>
<td><strong>Expansion</strong></td>
</tr>
<tr>
<td>$x = \sum_{i=1}^{n} c_i v_i$</td>
<td>$f(t) = \sum_{i=0}^{\infty} c_i v_i(t)$ means:</td>
</tr>
<tr>
<td>$\begin{bmatrix} v_1 &amp; \ldots &amp; v_n \end{bmatrix} \mathbf{c}$</td>
<td>$\sum_{i=0}^{n} c_i v_i(t) \rightarrow f(t) \text{ for } n \rightarrow \infty$</td>
</tr>
</tbody>
</table>
We have studied an efficient algorithm – filtered back projection (FBP) – for computing the CT reconstruction.

And we have also seen that the reconstruction is somewhat sensitive to noise in the data.

- How can we further study this sensitivity to noise?
- How can we possibly reduce the influence of the noise?
- What consequence does that have for the reconstruction?

We need a mathematical tool that lets us perform a detailed study of these aspects: the singular value decomposition/expansion.

But before going into these details, we will start with a simple example from signal processing, to explain the basic idea.
Assume that we know the characteristics of the system, and that we have measured the noisy output signal \( g(t) \). Now we want to reconstruct the input signal \( f(t) \).

The mathematical (forward) model, assuming \( 2\pi \)-periodic signals:

\[
g(t) = \int_{-\pi}^{\pi} h(\tau - t) f(\tau) \, d\tau \quad \text{or} \quad g = h \ast f \quad \text{(convolution)}.
\]

Here, the function \( h(t) \) (called the “impulse response”) defines the system.
In this example, the input $f(t)$ is white noise, and the output $g(t)$ is filtered noise. The corresponding Fourier transforms are $\hat{f}(\omega)$ and $\hat{g}(\omega)$.

**Deconvolution**: reconstruct the input $f$ from the output $g = h \ast f$. 
Fourier Series of Periodic Functions

Our tool is the Fourier series of a $2\pi$-periodic function $f$:

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{int}, \quad i = \sqrt{-1},$$

with the Fourier coefficients

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-int} f(t) \, dt = \frac{1}{2\pi} \langle \psi_n, f \rangle, \quad \psi_n = e^{int}.$$

We can think of the functions $\psi_n$ as a convenient basis for analysing the behavior of periodic functions. Similarly:

$$g(t) = \sum_{n=-\infty}^{\infty} d_n \psi_n, \quad d_n = \langle \psi_n, g \rangle.$$
Due to the linearity, we have

\[ g = h \ast f = h \ast \left( \sum_{n=-\infty}^{\infty} c_n \psi_n \right) = \sum_{n=-\infty}^{\infty} c_n (h \ast \psi_n). \]

Hence, all we need to know is the system’s response to each basis function \( \psi_n = e^{int} \).

For the periodic systems we consider here, the convolution of \( h \) with \( \psi_n \) produces a scaled version of \( \psi_n \):

\[ h \ast \psi_n = \mu_n \psi_n, \quad \text{for all } n, \]

where \( \mu_n = \langle \psi_n, h \rangle \) (we do not prove this). Hence,

\[ g = \sum_{n=-\infty}^{\infty} c_n \mu_n \psi_n = \sum_{n=-\infty}^{\infty} d_n \psi_n \iff f = \sum_{n=-\infty}^{\infty} \frac{d_n}{\mu_n} \psi_n. \]

Deconvolution is transformed to a simple algebraic operation: \textit{division}.
Naive Reconstruction from Noisy Data

Top left: the input $f(t)$ and the noisy output $g(t)$ (noise not visible).

Bottom left: corresponding Fourier coefficient; note the “noise floor.”

Bottom right: the reconstructed Fourier coefficients $d_n/\mu_n$ are dominated by the noise for $n > 100$, and a naive reconstruction is useless.

Top right: the naive and useless reconstruction.
▷ Left: same as previous slide.
▷ Bottom right: let us keep the first $\pm 100$ coefficients only.
▷ Top right: comparison of $f(t)$ and the truncated reconstruction using $\pm 100$ terms in the Fourier expansion. It captures the general shape of $f(t)$. 

Per Christian Hansen
SVD Analysis
What We Have Learned So Far

- With the right choice of basis functions, we can turn a complicated problem into a simpler one.
- Here: the basis functions are the complex exponentials; deconvolution $\rightarrow$ division in Fourier domain.
- Inspection of the expansion coefficients reveals how and when the noise enters in the reconstruction.
- Here: the noise dominates the output’s Fourier coefficients for higher frequencies, while the low-frequency coefficients are ok.
- We can avoid most of the noise (but not all) by means of filtering, at the cost of losing some details.
- Here: we simply truncate the Fourier expansion for the reconstruction.

Let us apply the same idea to parallel-beam CT reconstruction!
The Radon Transform


- The radon transform $g = \mathcal{R}f$ In previous slides $g$ was called $p_\theta(s)$.
- The image: $f(\xi)$ with $\xi = (x, y) \in D$, the disk with radius 1.
- The sinogram (the data): $g(s, \theta)$ with $s \in [-1, 1]$ and $\theta \in [0^\circ, 360^\circ]$. 
Singular Values and Singular Functions

There exist scalars \( \mu_{mk} \) and functions \( u_{mk}(s, \theta) \) and \( v_{mk}(\xi) \) such that

\[
\mathcal{R} v_{mk} = \mu_{mk} u_{mk}, \quad m = 0, 1, 2, 3, \ldots \quad k = 0, 1, 2, \ldots, m.
\]

The scalars are called the singular values:

\[
\mu_{mk} = 2 \sqrt{\frac{\pi}{m+1}} \quad \text{with multiplicity} \ m+1.
\]

The corresponding left and right singular functions are

\[
u_{mk}(\xi) \propto r^{|m-2k|} \cdot P_0^{0,|m-2k|}(2r^2-1) \cdot e^{-i(m-k)\phi},
\]

where \( r = \|\xi\|_2 \), \( U_m \) is the Chebyshev polynomial of the second kind, \( P_0^{0,|m-2k|} \) is the Jacobi polynomial, and \( \nu = \frac{1}{2}(m+|m-2k|) \).

(The word “singular” is used in the sense “special” or “unique.”)
About the Singular Functions

From the previous slide:

\[ \mathcal{R} u_{mk} = \mu_{mk} \nu_{mk}, \quad m = 0, 1, 2, 3, \ldots \quad k = 0, 1, 2, \ldots, m. \]

- The functions \( u_{mk} \) are an orthonormal basis for \([-1, 1] \times [0^\circ, 360^\circ] \).
- The functions \( \nu_{mk} \) are an orthonormal basis for the unit disk \( D \).
- Both functions are complex – similar to the Fourier basis.

The expansions of \( f \) and \( g \) take the form

\[
f(\xi) = \sum_{m,k} \langle \nu_{mk}, f \rangle \nu_{mk}(\xi), \quad g(s, \theta) = \sum_{m,k} \langle u_{mk}, g \rangle u_{mk}(s, \theta).
\]

\[
\langle \nu_{mk}, f \rangle = \int_D \overline{\nu_{mk}(\xi)} f(\xi) \, d\xi,
\]

\[
\langle u_{mk}, g \rangle = \int_{-1}^{1} \int_{0^\circ}^{360^\circ} \overline{u_{mk}(s, \theta)} g(s, \theta) \, d\theta \, ds.
\]
Some Left Singular Functions for $m = 0, 1, 2, 3, 4$
Some Right Singular Functions for $m = 0, 1, 2, 3, 4$
Singular Values, and Some Comments

- All singular values are positive (no zeros); they **decay** rather slowly.
- If $\mu_{mk} = \mu_j$ with $j = \frac{1}{2} m(m+1) + k + 1$, then $\mu_j \propto 1/\sqrt{j}$ for large $j$.
- Singular functions with higher index $j$ have higher frequencies.
- The higher the frequency, the more the damping in $\mathcal{R} v_{mk} = \mu_{mk} u_{mk}$.
- Hence the Radon transform $g = \mathcal{R} f$ is a “smoothing” operation
- ... and the reverse operation $f = \mathcal{R}^{-1} g$ amplifies higher frequencies!

These are intrinsic properties of the mathematical problem itself.
These are the coefficients $\langle u_{mk}, g \rangle$ for the sinogram corresponding to the Shepp-Logan phantom – ordered according to increasing index $m$.

They decay, as expected. The specific behavior for $k = 0, \ldots, m$ is due to the symmetry of the phantom.
The Inverse Radon Transform is Unbounded

From *linear algebra* we know that if \( b = Ax \iff x = A^{-1}b \) then

\[
\|b\|_2 \leq \|A\| \|x\|_2 \quad \text{and} \quad \|x\|_2 \leq \|A^{-1}\| \|b\|_2, \quad \|A\|^2 = \sum_{i,j} a_{ij}^2.
\]

Hence, a perturbation \( \Delta b \) of \( b \) produces a reconstruction perturbation \( \Delta x = A^{-1}\Delta b \) with \( \|\Delta x\|_2 \leq \|A^{-1}\| \|\Delta b\|_2 \).

For the *Radon transform* \( g = \mathcal{R}f \iff f = \mathcal{R}^{-1}g \) we have

\[
\|g\|_2 \leq \|\mathcal{R}\| \|f\|_2 \quad \text{with} \quad \|\mathcal{R}\|^2 = \sum_{m,k} \mu_{mk}^2
\]

\[
\|f\|_2 \leq \|\mathcal{R}^{-1}\| \|g\|_2 \quad \text{with} \quad \|\mathcal{R}^{-1}\|^2 = \sum_{m,k} \frac{1}{\mu_{mk}^2}.
\]

Trouble: \( \|\mathcal{R}^{-1}\| = \infty \). An unperturbed \( g = \mathcal{R}f \) satisfies the *Picard condition*. But a noisy \( g \) does not, and noisy reconstruction is unbounded!
Due to the factor $\sqrt{1 - s^2}$, all the left singular functions satisfy

$$u_{mk}(s, \theta) \to 0 \quad \text{for} \quad s \to \pm 1.$$ 

This reflects the fact that rays through $D$ that almost grace the edge of the disk contribute very little to the sinogram.

This puts a restriction on sinograms $g(s, \theta)$ that admit a reconstruction:

- The sinogram $g = \mathcal{R} f$ is a sum of the singular functions $u_{mk}$.
- Hence, the sinogram inherits the property $g(s, \theta) \to 0$ for $s \to \pm 1$.
- A perturbation $\Delta g$ of $g$ that does not have this property may not produce a bounded perturbation $\mathcal{R}^{-1} \Delta g$ of $f$. 
Illustration of the “Range Condition”

When the noise violates the “range condition” $g(s, \theta) \to 0$ for $s \to \pm 1$.

- Top row: the noise in the sinogram maps to large errors in the corners of the reconstructed image.
- Note that this phenomenon is not restricted to FBP and disk domains!
- Fix the problem: add damping to the sinogram data near $s = \pm 1$. 
Let’s Reconstruct

In terms of the singular values and functions, the inverse Radon transform takes the form

\[ f(\xi) = \sum_{m,k} \frac{\langle u_{mk}, g \rangle}{\mu_{mk}} v_{mk}(\xi). \]

Since the image \( f(\xi) \) has finite norm (finite energy), we conclude that the magnitude of the coefficient \( \frac{\langle u_{mk}, g \rangle}{\mu_{mk}} \) must decay “sufficiently fast.”

**The Picard Condition.** The expansion coefficients \( \langle u_{mk}, g \rangle \) for \( g(s, \theta) \) must decay sufficiently faster than the singular values \( \mu_{mk} \), such that

\[ \sum_{m,k} \left| \frac{\langle u_{mk}, g \rangle}{\mu_{mk}} \right|^2 < \infty. \]

When noise is present in the sinogram \( g(s, \theta) \), then this condition is not satisfied for large \( m \) (cf. the signal restoration example from before). This calls for some kind of filtering.
Let’s Introduce Filters

A simple remedy for the noise-magnification, by the division with $\mu_{mk}$, is to introduce filtering:

$$f(\xi) = \sum_{m,k} \varphi_{mk} \frac{\langle u_{mk}, g \rangle}{\mu_{mk}} v_{mk}(\xi).$$

The filter factors $\varphi_{mk}$ must decay fast enough that they, for large $m$, can counteract the factor $\mu_{mk}^{-1}$. More on this later in the course.

We can think of the filter factors as *modifiers* of the expansion coefficients $\langle u_{mk}, g \rangle$ for the sinogram.

In other words, they ensure that the filtered coefficients $\varphi_{mk} \langle u_{mk}, g \rangle$ decay fast enough to satisfy the Picard condition from the previous slide.

*The filtering inevitably dampens the higher frequencies associated with the small $\mu_{mk}$, and hence some details and edges in the image are lost.*
Recall the filtered back projection (FBP) algorithm:

1. For fixed $\theta$ compute the Fourier transform $\hat{g}(\omega, \theta) = \mathcal{F}(g(s, \theta))$.
2. Apply the ramp filter $|\omega|$ and compute the inverse Fourier transform $g_{\text{filt}}(s, \theta) = \mathcal{F}^{-1}(|\omega| \hat{g}(\omega, \theta))$.
3. Do the above for all $\theta \in [0^\circ, 360^\circ]$.
4. Then compute $f(\xi) = \int_{0^\circ}^{360^\circ} g_{\text{filt}}(x \cos \theta + y \sin \theta, \theta) \, d\theta$, $\xi = (x, y)$.

It is the ramp filter $|\omega|$ in step 2 that magnifies the higher frequencies in the sinogram $g(s, \theta)$.

This amplification is \textit{equivalent} to the division by the singular values $\mu_{mk}$ in the above analysis.
Filtered Back Projection, now with Low-Pass Filtering

How the filtered back projection algorithm (FBP) is really implemented:

1. Choose a low-pass filter $\varphi_{LP}(\omega)$.
2. For every $\theta$ compute the Fourier transform $\hat{g}(\omega, \theta) = \mathcal{F}(g(s, \theta))$.
3. Apply the combined ramp & low-pass filter, and compute the inverse Fourier transform $g_{filtfilt}(s, \theta) = \mathcal{F}^{-1}(|\omega| \varphi_{LP}(\omega) \hat{g}(\omega, \theta))$.
4. Then $f_{rec}(\xi) = \int_{0}^{360^\circ} g_{filtfilt}(x \cos \theta + y \sin \theta, \theta) \, d\theta$.

The low-pass filter $\varphi_{LP}(\omega)$ counteracts the ramp filter $|\omega|$ for large $\omega$. It is equivalent to the filter factors $\varphi_{mk}$ introduced on slide 23.
A couple of interesting quotes:

- “Theorem 4.2. A finite set of radiographs tells nothing at all.”
  

- “This is contrary [...] to the experience obtained in practice, where CT scanners produce reliable pictures [...] although only a finite number of projections are used.”
  
Vectors and Images/Sinograms

Note that in the discretized problem $A x = b$, the image and the sinogram are represented by the vectors $x$ and $b$, respectively.

While this is a convenient notation when we need the language of linear algebra, they are really 2D arrays and they should be visualized as such.

In Matlab notation:

```matlab
imagesc( reshape(x,N,N) ), axis image
imagesc( reshape(b,N_s,N_θ) ), axis image
```

where $N_s = \text{number of detector pixels}$, and $N_θ = \text{number of projections}$.

Going from an image $X$ to a vector $x$ is simple: just write $x = X(:)$. 
Matrix Notation and Interpretation

Notation:

\[
A = \begin{pmatrix}
\mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n
\end{pmatrix} = \begin{pmatrix}
\mathbf{r}_1 \\
\vdots \\
\mathbf{r}_m
\end{pmatrix}.
\]

The matrix \( A \) maps the discretized absorption coefficients (the vector \( x \)) to the data in the detector pixels (the elements of the vector \( b \)) via:

\[
b = \begin{pmatrix}
b_1 \\
b_2 \\
\vdots \\
b_m
\end{pmatrix} = A x = \begin{pmatrix}
\mathbf{x}_1 \mathbf{c}_1 + \mathbf{x}_2 \mathbf{c}_2 + \cdots + \mathbf{x}_n \mathbf{c}_n
\end{pmatrix} = \begin{pmatrix}
\mathbf{r}_1 \cdot x \\
\mathbf{r}_2 \cdot x \\
\vdots \\
\mathbf{r}_m \cdot x
\end{pmatrix}.
\]

See next slides for examples of the column and row interpretations.
Example of Column Interpretation

A $32 \times 32$ image has four nonzero pixels with intensities 1, 0.8, 0.6, 0.4. In the vector $\mathbf{x}$, these four pixels correspond to entries 468, 618, 206, 793. Hence the sinogram, represented as a vector $\mathbf{b}$, takes the form

$$
\mathbf{b} = 0.6 \mathbf{c}_{206} + 1.0 \mathbf{c}_{468} + 0.8 \mathbf{c}_{618} + 0.4 \mathbf{c}_{793}.
$$

Note that each pixel is mapped to a single sinusoidal curve in the sinogram.
Example of Row Interpretation

The $i$th row of $A$ maps $x$ to detector element $i$ via the $i$th ray:

$$b_i = r_i \cdot x = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, 2, \ldots, m.$$ 

This inner product approximates the line integral along ray $i$ in the Radon transform.

A small example:

$$a_{ij} = \text{length of ray } i \text{ in pixel } j$$

$$r_i = \begin{bmatrix} a_{i1} & a_{i2} & 0 & a_{i4} & 0 & 0 & a_{i7} & 0 & 0 \end{bmatrix}$$

$$b_i = r_i \cdot x = a_{i1}x + a_{i2}y + a_{i4}x_4 + a_{i7}x_7$$

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>4</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>5</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>
Recall the back projection,

\[ f(\xi) = \int_0^{2\pi} g(x \cos \theta + y \sin \theta, \theta) \, d\theta, \]

where we integrate \( g \) along a sinusoidal curve in the sinogram:

\[ x = A^T b = \begin{pmatrix} | & | & | \\ c_1 & \cdots & c_n \end{pmatrix}^T b = \begin{pmatrix} c_1^T \\ \vdots \\ c_n^T \end{pmatrix} b = \begin{pmatrix} c_1 \cdot b \\ \vdots \\ c_n \cdot b \end{pmatrix} \]

where each inner product \( c_j \cdot b \) corresponds to the above integration.
Back Projection – Alternative Interpretation

Another formula for back projection:

\[ x = A^T b = \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}^T b = \begin{pmatrix} | & | & \cdots & | \\ r_1 & \cdots & r_m \end{pmatrix} b \]

The interpretation here is that each data \( b_i \) is “distributed back” along the \( i \)th ray to the pixels, with “weight” = \( a_{ij} \):

\[ x = \sum_{i=1}^{m} b_i r_i = \sum_{i=1}^{m} \begin{pmatrix} b_i a_{i1} \\ \vdots \\ b_i a_{in} \end{pmatrix}. \]

The small example from before:

\[ b_i r_i = [ b_i a_{i1} \ b_i a_{i2} \ 0 \ b_i a_{i4} \ 0 \ 0 \ b_i a_{i7} \ 0 \ 0 ] \]
And Now: The Singular Value Decomposition

Assume that $A$ is $m \times n$ and, for simplicity, also that $m \geq n$:

\[ A = U \Sigma V^T = \sum_{i=1}^{n} u_i \sigma_i v_i^T. \]

Here, $\Sigma$ is a diagonal matrix with the *singular values*, satisfying

\[ \Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n), \quad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0. \]

The matrices $U$ and $V$ consist of *singular vectors*

\[ U = (u_1, \ldots, u_n), \quad V = (v_1, \ldots, v_n) \]

and both matrices have orthonormal columns: $U^T U = V^T V = I_n$.

Then $\|A\|_2 = \sigma_1$, $\|A^{-1}\|_2 = \|V \Sigma^{-1} U^T\|_2 = \sigma_n^{-1}$, and

\[ \text{cond}(A) = \frac{\|A\|_2 \|A^{-1}\|_2 = \frac{\sigma_1}{\sigma_n}}{.} \]
The matrix $\mathbf{A}^T \mathbf{A}$ is symmetric and hence 1) its eigenvalues are real, and 2) its eigenvectors are real and orthonormal (standard linear algebra stuff):

$$\mathbf{A}^T \mathbf{A} \mathbf{v}_i = \sigma_i^2 \mathbf{v}_i, \quad i = 1, 2, \ldots, n.$$ 

I.e., the right singular vectors $\mathbf{v}_i$ are the eigenvectors of $\mathbf{A}^T \mathbf{A}$ and the squared singular values are the corresponding eigenvalues.

This is not how the SVD should be computed – use only good numerical software. In Matlab:

- Use $[\mathbf{U}, \mathbf{S}, \mathbf{V}] = \text{svd}(\mathbf{A})$ or $[\mathbf{U}, \mathbf{S}, \mathbf{V}] = \text{svd}(\mathbf{A}, 0)$ to compute the full or “economy-size” SVD.
- Use $[\mathbf{U}, \mathbf{S}, \mathbf{V}] = \text{svds}(\mathbf{A})$ to efficiently compute a partial SVD (the largest singular values and corresponding singular vectors).
Important SVD Relations

Relations similar to the analysis of the Radon transform:

\[ A v_i = \sigma_i u_i, \quad \| A v_i \|_2 = \sigma_i, \quad i = 1, \ldots, n. \]

Also, if \( A \) is nonsingular, then

\[ A^{-1} u_i = \sigma_i^{-1} v_i, \quad \| A^{-1} u_i \|_2 = \sigma_i^{-1}, \quad i = 1, \ldots, n. \]

These equations are related to the solution:

\[
\begin{align*}
x &= \sum_{i=1}^{n} (v_i \cdot x) v_i \\
A x &= \sum_{i=1}^{n} \sigma_i (v_i \cdot x) u_i, \quad b = \sum_{i=1}^{n} (u_i \cdot b) u_i \\
A^{-1} b &= \sum_{i=1}^{n} \frac{u_i \cdot b}{\sigma_i} v_i.
\end{align*}
\]
The singular values decay to zero, with no gap in the spectrum. The decay rate determines how difficult the problem is to solve. Discretization of the Radon transformation gives a mildly ill-posed problem.
The singular vectors (1D problem) have increasing frequency as $i$ increases.
“Picard Plot” – Without and With Noise

Left: no noise. Singular values $\sigma_i$ and rhs coefficients $|u_i^Tb|$ both level off at the machine precision.

Right: with noise. Now the rhs coefficients $|u_i^Tb|$ level off at the noise level, and only $\approx 18$ SVD components are reliable.
The “Naive” Solution

From now on, the coefficient matrix $A$ is allowed to have more rows than columns, i.e.,

$$A \in \mathbb{R}^{m \times n} \quad \text{with} \quad m \geq n.$$ 

For $m > n$ it is natural to consider the least squares problem

$$\min_x \|Ax - b\|_2.$$ 

When we say “the naive solution” $x^\text{naive}$ we either mean the solution $A^{-1}b$ (when $m = n$) or the least squares solution (when $m > n$).

We emphasize the convenient fact that the naive solution has precisely the same SVD expansion in both cases:

$$x^\text{naive} = \sum_{i=1}^{n} \frac{u_i \cdot b}{\sigma_i} v_i.$$ 

Per Christian Hansen

SVD Analysis
Ill-Conditioned Problems

Discrete ill-posed problems are characterized by having coefficient matrices with a large condition number. The solution is very sensitive to errors in the data.

Specifically, assume that the exact and perturbed solutions $\bar{x}$ and $x$ satisfy

$$A\bar{x} = \bar{b}, \quad Ax = b = \bar{b} + e,$$

where $e$ denotes the perturbation (the errors and noise). Then classical perturbation theory leads to the bound

$$\frac{\|\bar{x} - x\|_2}{\|\bar{x}\|_2} \leq \text{cond}(A) \frac{\|e\|_2}{\|\bar{b}\|_2}.$$

Since $\text{cond}(A) = \sigma_1/\sigma_n$ is large, this implies that $x^{\text{naive}} = A^{-1}b$ can be very far from $\bar{x}$.

Per Christian Hansen

SVD Analysis
The Geometry of Ill-Conditioned Problems

\[ \mathbb{R}^n = \text{span}\{v_1, \ldots, v_n\} \]

\[ \mathbb{R}^m = \text{span}\{u_1, \ldots, u_m\} \]

Exact sol.: \( \bar{x} \)

Naive sol.: \( x^{\text{naive}} \)

\[ \bar{b} = A \bar{x} \]

\[ b = \bar{b} + e \]
The Need for Filtering

Recall that the (least squares) solution is given by

\[
x^{\text{naive}} = \sum_{i=1}^{n} \frac{u_i \cdot b}{\sigma_i} v_i.
\]

When noise is present in the data \( b = \bar{b} + e \), then

\[
u_i \cdot b = u_i \cdot \bar{b} + u_i \cdot e \approx \begin{cases} u_i \cdot \bar{b}, & |u_i \cdot \bar{b}| > |u_i \cdot e| \\
u_i \cdot e, & |u_i \cdot \bar{b}| < |u_i \cdot e|.
\end{cases}
\]

We note that:

- Due to the Picard condition, the noise-free \( |u_i \cdot \bar{b}| \) decay.
- The “noisy” components \( |u_i \cdot b| \) are those for which \( |u_i \cdot \bar{b}| \) is small,
- and they correspond to the smaller singular values \( \sigma_i \).
Spectral Filtering

Many of the noise-reducing methods treated in this course produce solutions which can be expressed as a filtered SVD expansion of the form

$$x_{\text{filt}} = \sum_{i=1}^{n} \varphi_i \frac{u_i \cdot b}{\sigma_i} v_i,$$

where $\varphi_i$ are the filter factors associated with the method.

These methods are called spectral filtering methods because the SVD basis can be considered as a spectral basis.

A simple approach is to discard the SVD coefficients corresponding to the smallest singular values:

$$\varphi_i^{\text{TSVD}} = \begin{cases} 1 & i = 1, 2, \ldots, k \\ 0 & \text{else.} \end{cases}$$

More sophisticated methods will be discussed in the third week.
Define the truncated SVD (TSVD) solution as

\[
\mathbf{x}_k = \sum_{i=1}^{n} \varphi_i^{TSVD} \frac{u_i \cdot \mathbf{b}}{\sigma_i} \mathbf{v}_i = \sum_{i=1}^{k} \frac{u_i \cdot \mathbf{b}}{\sigma_i} \mathbf{v}_i, \quad k < n.
\]

We can show that if \( \text{Cov}(\mathbf{b}) = \eta^2 \mathbf{I}_m \) then

\[
\text{Cov}(\mathbf{x}_k) = \eta^2 \sum_{i=1}^{k} \frac{1}{\sigma_i^2} \mathbf{v}_i \mathbf{v}_i^T
\]

and therefore

\[
\|\mathbf{x}_k\|_2 \ll \|\mathbf{x}^\text{naive}\|_2 \quad \text{and} \quad \|\text{Cov}(\mathbf{x}_k)\|_2 \ll \|\text{Cov}(\mathbf{x}^\text{naive})\|_2.
\]

The prize we pay for smaller covariance is bias: \( \mathcal{E}(\mathbf{x}_k) \neq \mathcal{E}(\mathbf{x}^\text{naive}). \)
Theorem

Let $\mathbf{b} = \bar{\mathbf{b}} + \mathbf{e}$ and let $\mathbf{x}_k$ and $\bar{\mathbf{x}}_k$ denote the TSVD solutions computed with the same $k$.

Then

$$\frac{\|\bar{\mathbf{x}}_k - \mathbf{x}_k\|_2^2}{\|\mathbf{x}_k\|_2^2} \leq \frac{\sigma_1}{\sigma_k} \frac{\|\mathbf{e}\|_2}{\|\mathbf{A}\mathbf{x}_k\|_2}.$$

We see that the condition number for the TSVD solution is

$$\kappa_k = \frac{\sigma_1}{\sigma_k}$$

and it can be much smaller than $\text{cond}(\mathbf{A}) = \sigma_1/\sigma_n$. 
As we increase the truncation parameter in the TSVD solution $x_k$, we include more SVD components and also more noise. At some point the noise becomes visible and then starts to dominate $x_k$. 
The Truncation Parameter

Note: the truncation parameter $k$ in

$$x_k = \sum_{i=1}^{k} \frac{u_i \cdot b}{\sigma_i} v_i$$

is dictated by the coefficients $u_i^T b$, not the singular values.

Basically we should choose $k$ as the index $i$ where $|u_i \cdot b|$ start to “level off” due to the noise.

The TSVD solution and residual norms vary monotonically with $k$:

$$\|x_k\|_2^2 = \sum_{i=1}^{k} \left( \frac{u_i \cdot b}{\sigma_i} \right)^2 \leq \|x_{k+1}\|_2^2,$$

$$\|A x_k - b\|_2^2 = \sum_{i=k+1}^{n} (u_i \cdot b)^2 \geq \|A x_{k+1} - b\|_2^2.$$
Where TSVD Fits in the Picture

\[ \mathbb{R}^n = \text{span}\{v_1, \ldots, v_n\} \quad \text{and} \quad \mathbb{R}^m = \text{span}\{u_1, \ldots, u_m\} \]

**Exact sol.**: \( \bar{x} \)

**Naive sol.**: \( x^{\text{naive}} \)

\[ \bar{b} = A \bar{x} \quad \text{and} \quad b = \bar{b} + e \]

\( R. C. Hansen \)

SVD Analysis