Inverse Problems (an introduction oriented to data)

MINICOURSE IPMAS 2022

AXEL OSSES (CMM. U. DE CHILE),

CARLOS CASTILLO (PUC)

JANUARY 12-14 2022

An inverse problem is like solving an enigma

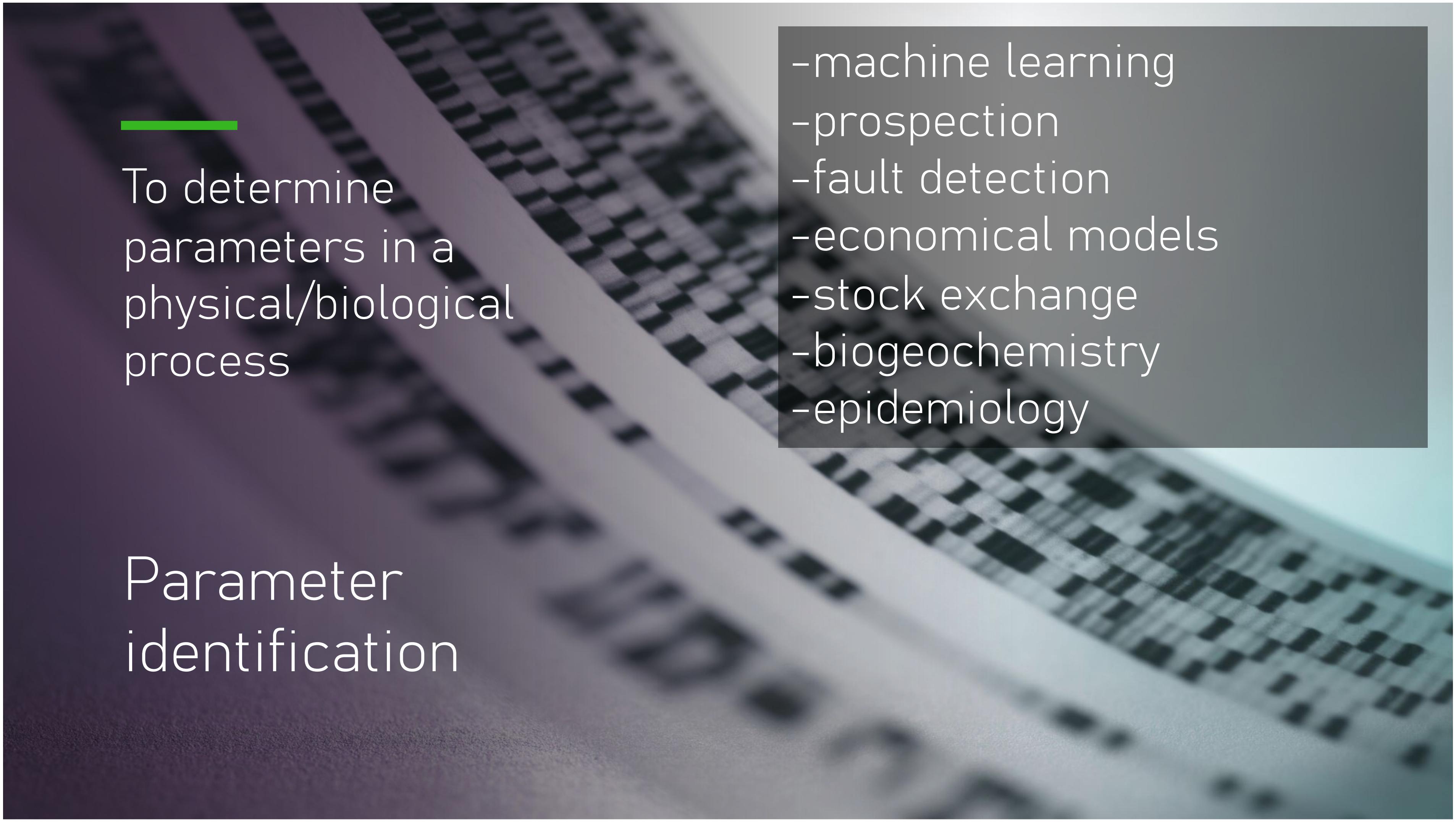
" The best recipe for the detective novel: the detective must never know more than the reader" Agatha Christie

A → B	Cause → Effect	DIRECT
B → A	Effect → Cause	INVERSE

Combining models and data in the best possible way

- weather forecast
- atmospheric chemistry
- seismicity
- oceanography
- satellites, remote sensing
- geolocalization
- hidrology

Data Assimilation


To determine parameters in a physical/biological process
Parameter identification
-machine learning
-prospection
-fault detection
-economical models
-stock exchange
-biogeochemistry
-epidemiology

Extracting information from an image or synthesising an image

Image processing

-interferometría
-tomografía
-biomedicina
-

-compression
-reconstruction
-synthesis
-segmentation
-classification
-detection
-registration
-deconvolution

From the effects to determine the causes......

$$A \rightarrow B = f(A)$$

$$B \rightarrow A = f^{-1}(B)$$

Causes → Effects DIRECT

Effects → Causes INVERSE

Inverse Problems

Inverse Problem	Cause A	Effect B	Function A = f(B)
Medical diagnosis	Disease	Symptoms	Physiology, genetics
Crime in a detective novel	Mobile	Murder	Human psychology
Interferometry	Intensity emission	Radio signal	Fourier transform
MRI	Density of living tissues	Magnetization	Bloch's equations, Fourier transform
Example 1 → **Tomography**	Atenuación del medio	Sombras por rayos X	Radon transform
Example 2 → **Deblurring**	Original image	Blurred image	Convolution
Electrophysiology	Currents	Voltages	Laplace equation
Seismology	Source	Displacement, velocity, acceleration	Wave equation, elasticity

Two examples of inverse problems:

TEHORETICAL BASIS... THIS PRESENTATION

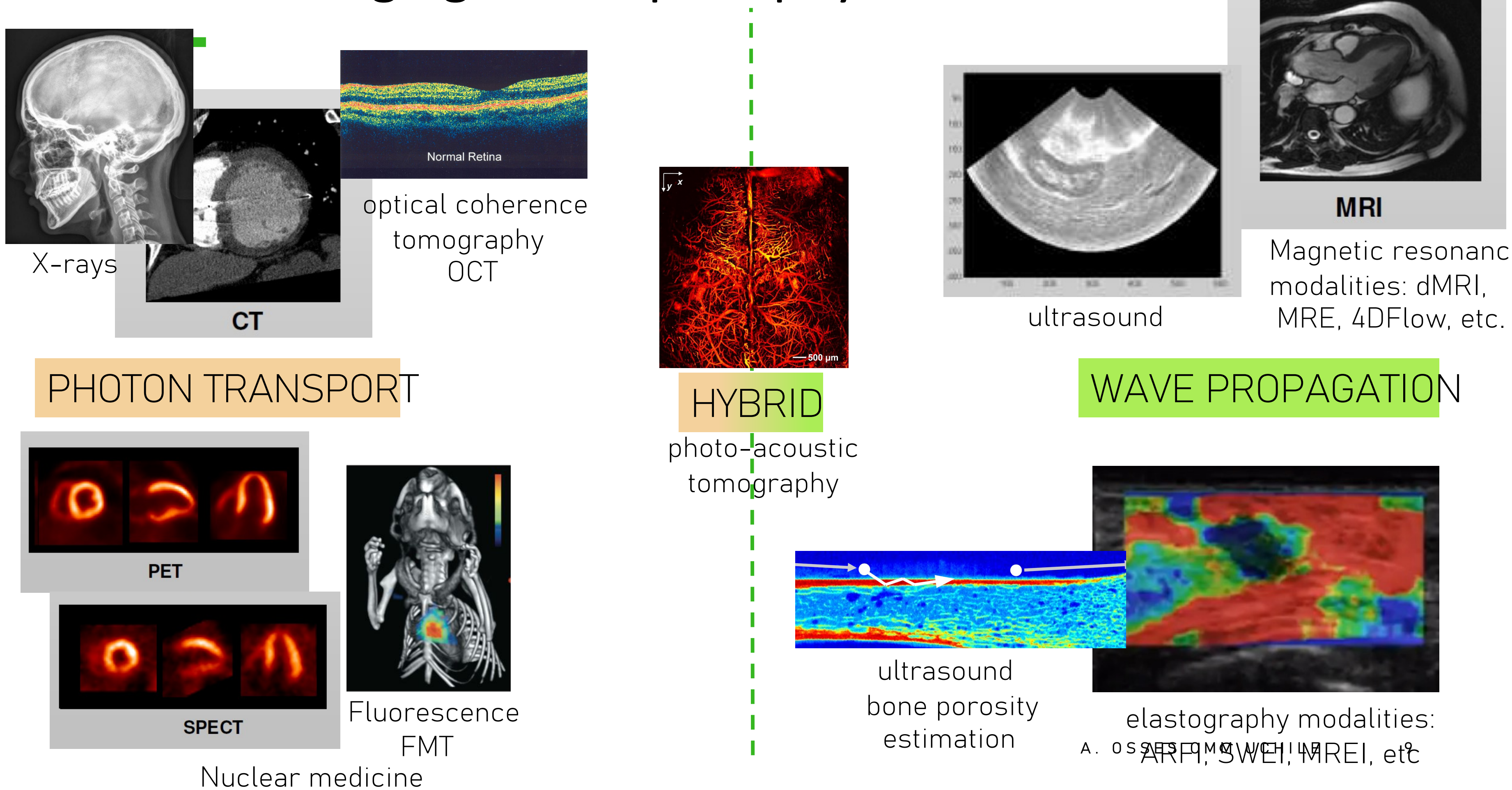
IN PRACTICE... SESSIONS WITH JUPYTER NOTEBOOKS AND PYTHON

1. Radon transform and X-ray tomography

2. Bringing a blurred image into focus using its singular value representation

Radon transform and X-ray tomography.

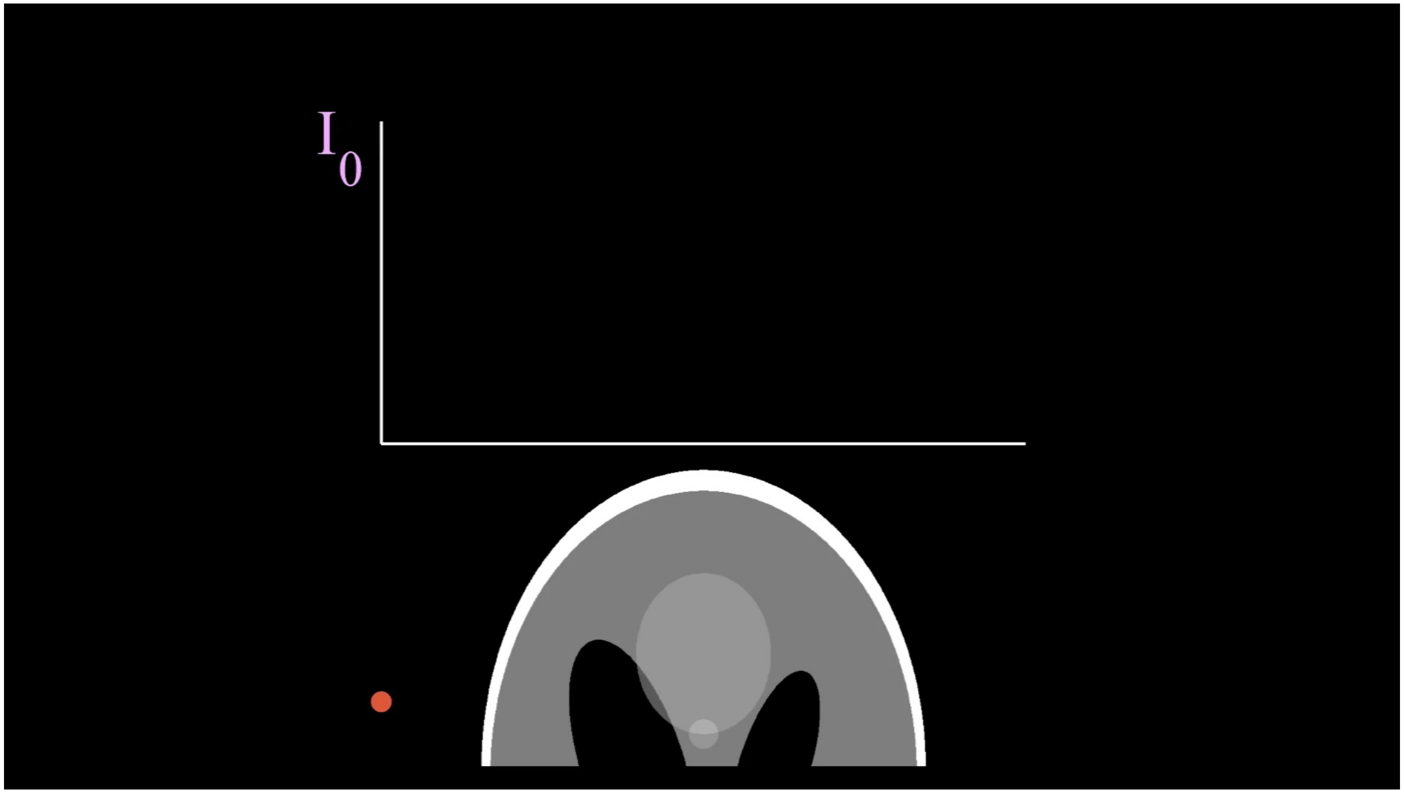
biomedical imaging techniques: physics



X-ray attenuation

credit: Samuli Siltanen

the measurent $\log(I_0/I_1)$ is the integral of the attenuation along the line



The Shepp-Logan phantom: Larry Shepp and Benjamin F. Logan for their 1974 paper *The Fourier Reconstruction of a Hea*

Beer's law

(monochromatic, X-ray beam
no refraction or diffraction)

Intensity decay...

$$\Delta I = I_{out} - I_{in}$$

is proportional to attenuation, thickness and intensity...

$$\Delta I = -a(x) I \Delta x$$

The infinitesimal change...

$$\frac{dI}{dx} = -a(x) I$$

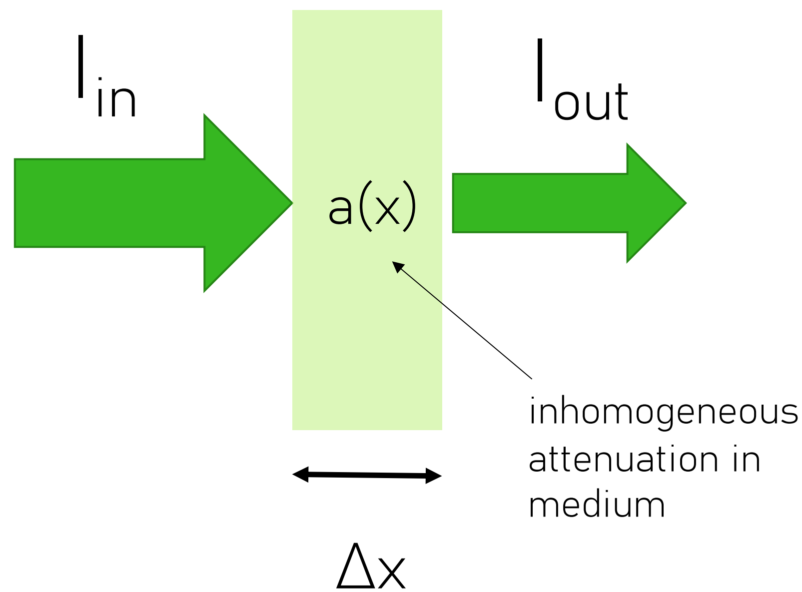
gives by integration on lines the solution:

$$I = I_0 \exp\left(-\int_L a\right)$$

$$L : \text{line from } x_0 \text{ to } x$$

so the sum over lines can be measured

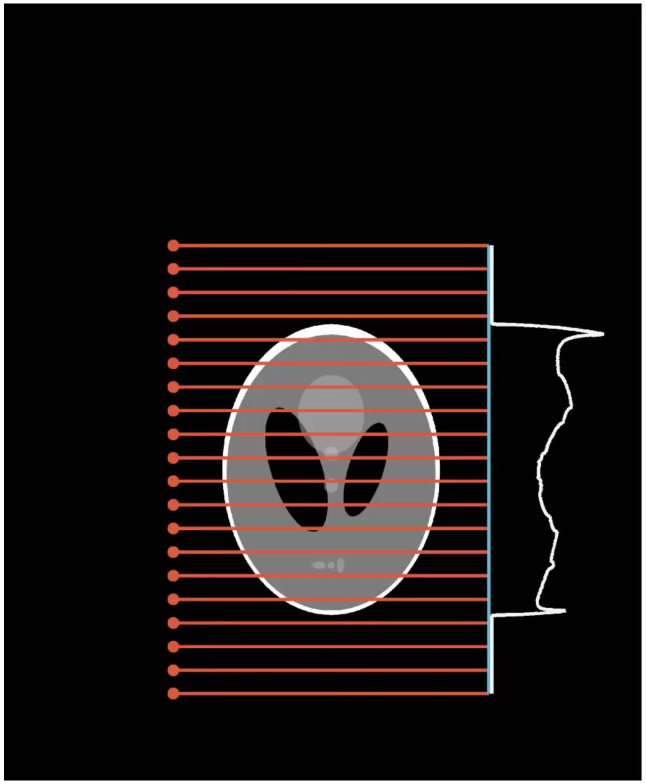
$$\int_L a = -\ln \frac{I}{I_0}$$



CT - scanner

image? → sinogram (Radon transform)

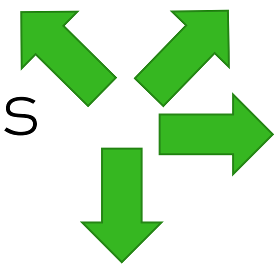
Mathematical tool:
the inverse Radon transform



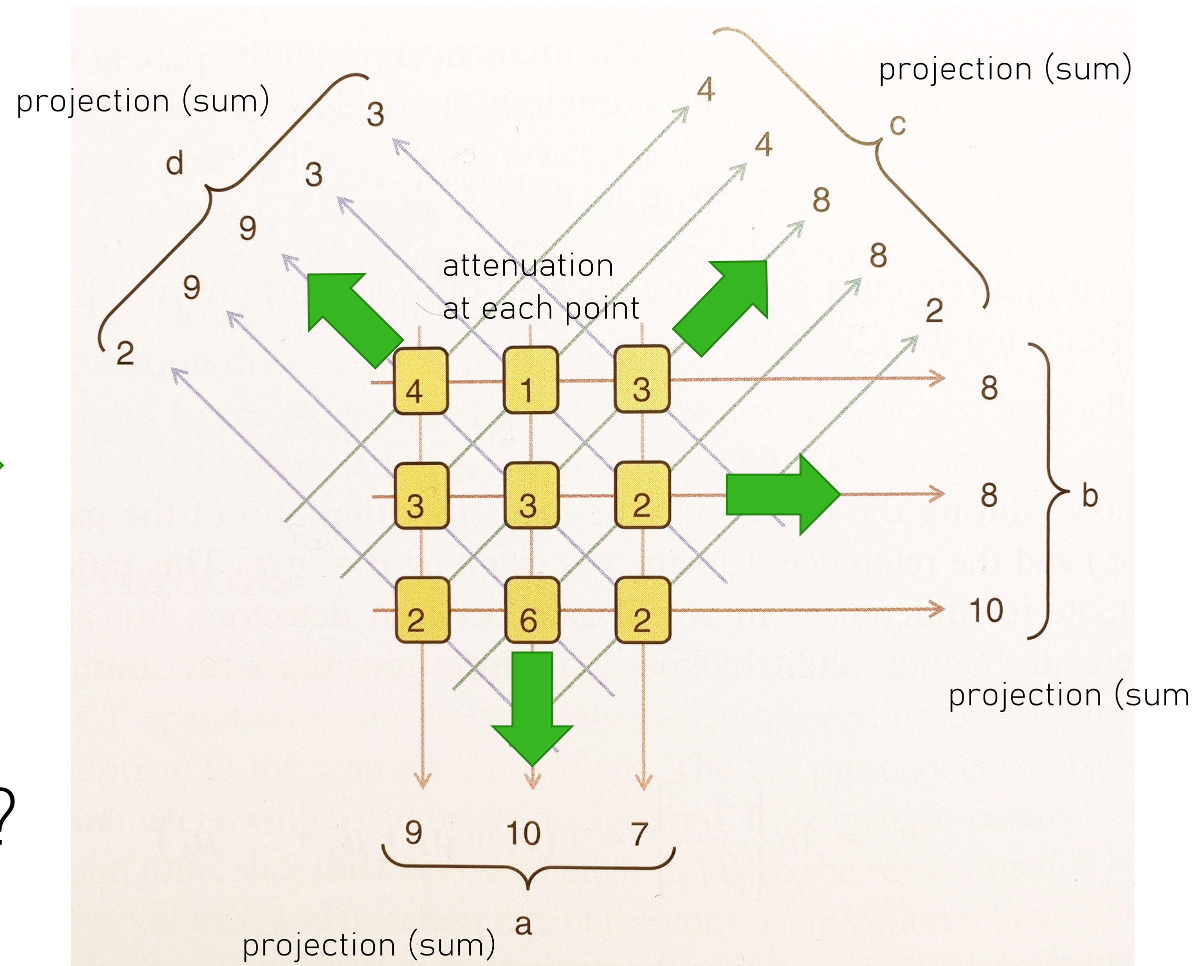
credit: Samuli Siltanen

Principle of Computed Tomography:

Projections

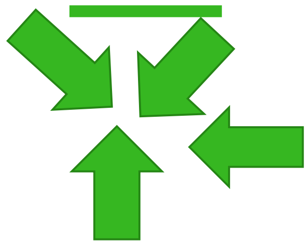


How to recover the attenuation from sum of lines?



credits figure: ref 2

Backprojections

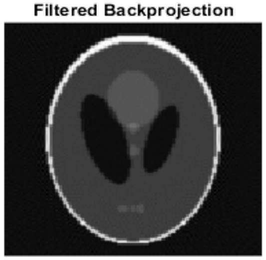


30	25	26
30	35	26
29	37	28

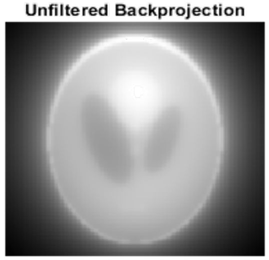
*9/26/4

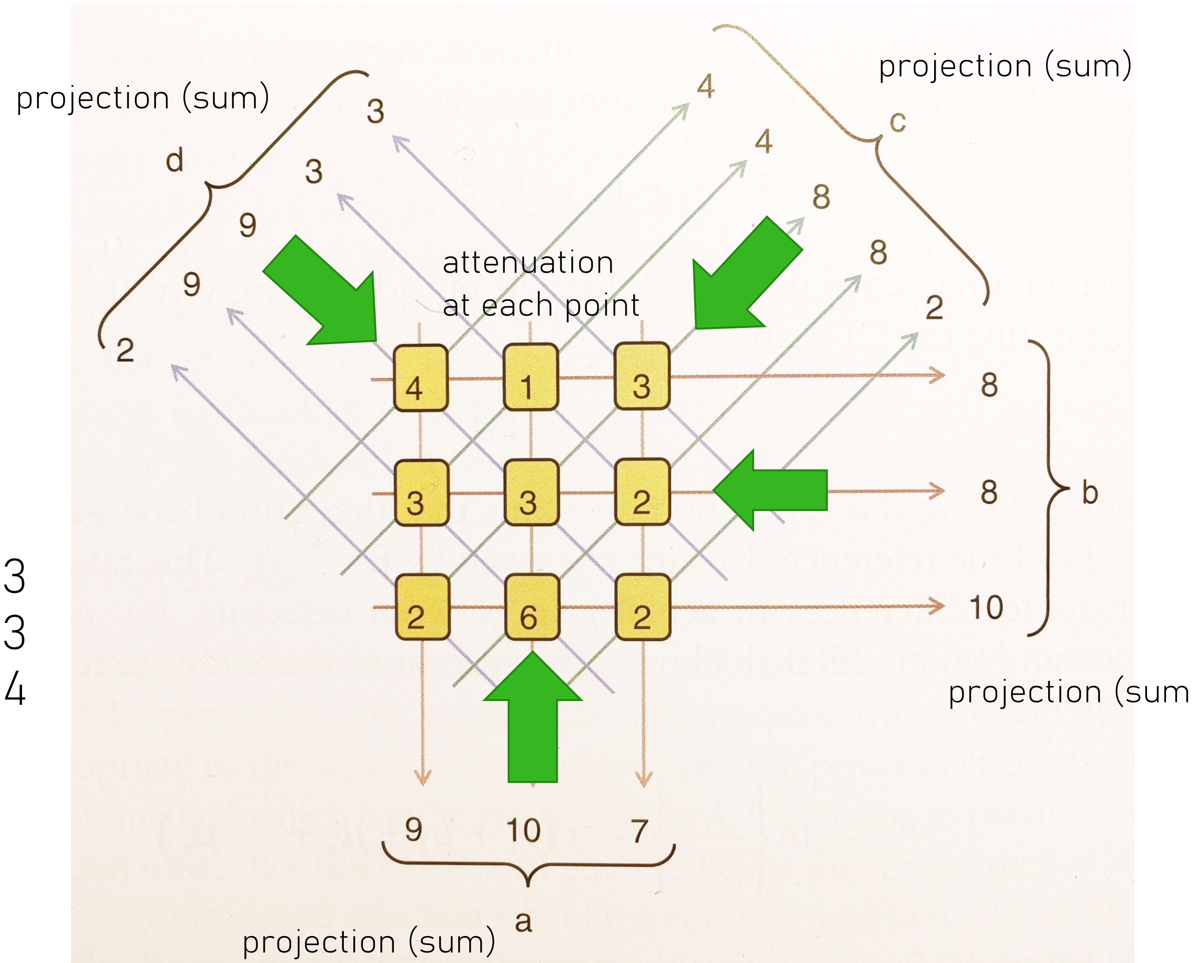
2.6	2.2	2.3
2.6	3.0	2.3
2.5	3.2	2.4

Filtered Backprojection



Unfiltered Backprojection





credits figure: ref 2

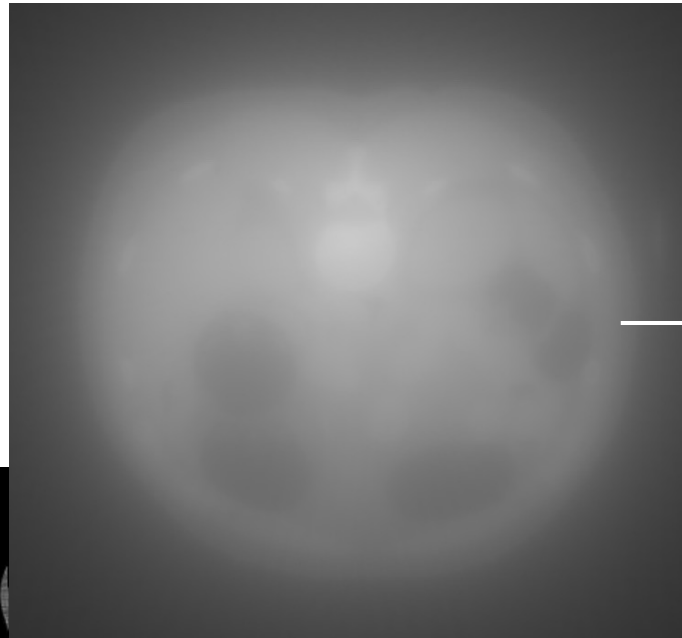
Backprojection:
effect of an increasing
number of projections

original (axial view)

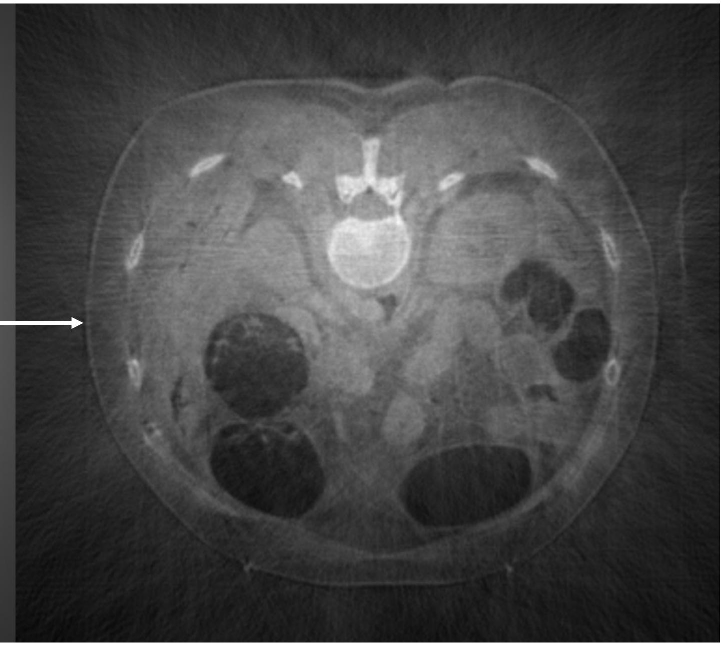
magic!

Deconvolution Filter

original (axial view)

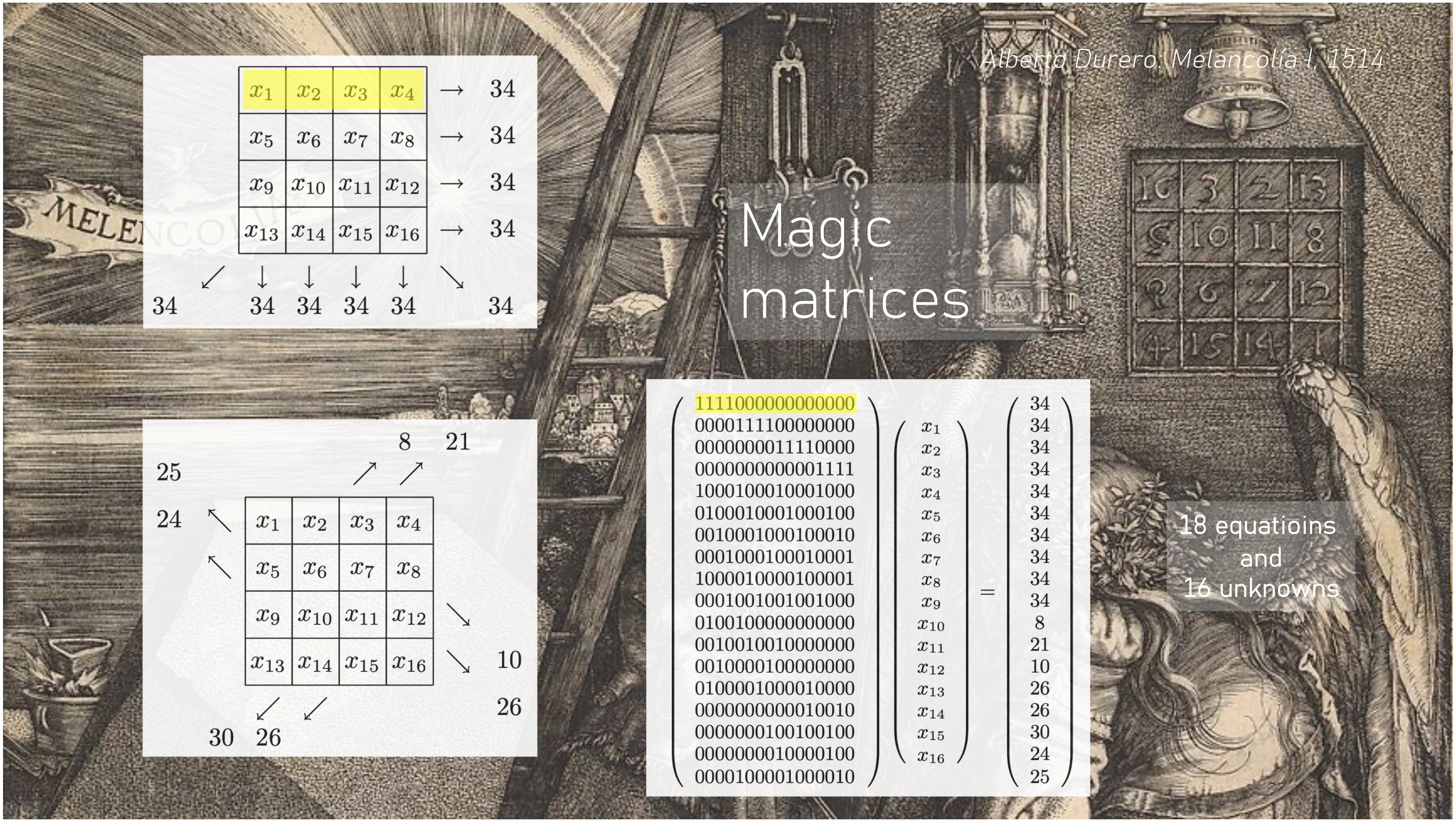


backprojection



filtered backprojection

Source:
http://www.perlproductions.at/index.php?choice=referenz&lang=en&id=15



Magic matrices

	x_1	x_2	x_3	x_4	→ 34
	x_5	x_6	x_7	x_8	→ 34
	x_9	x_{10}	x_{11}	x_{12}	→ 34
	x_{13}	x_{14}	x_{15}	x_{16}	→ 34
↙ 34	↓ 34	↓ 34	↓ 34	↓ 34	↘ 34

			↗ 8	↗ 21	
25 ↖	x_1	x_2	x_3	x_4	
24 ↖	x_5	x_6	x_7	x_8	
	x_9	x_{10}	x_{11}	x_{12}	↘ 10
	x_{13}	x_{14}	x_{15}	x_{16}	↘ 26
	↙ 30	↙ 26			

$$
\begin{pmatrix}
1111000000000000\\
0000111100000000\\
0000000011110000\\
0000000000001111\\
1000100010001000\\
0100010001000100\\
0010001000100010\\
0001000100010001\\
1000010000100001\\
0001001001001000\\
0100100000000000\\
0010010010000000\\
0010000100000000\\
0100001000010000\\
0000000000010010\\
0000000100100100\\
0000000010000100\\
0000100001000010
\end{pmatrix}
\begin{pmatrix}
x_1\\ x_2\\ x_3\\ x_4\\ x_5\\ x_6\\ x_7\\ x_8\\ x_9\\ x_{10}\\ x_{11}\\ x_{12}\\ x_{13}\\ x_{14}\\ x_{15}\\ x_{16}
\end{pmatrix}
=
\begin{pmatrix}
34\\ 34\\ 34\\ 34\\ 34\\ 34\\ 34\\ 34\\ 34\\ 34\\ 8\\ 21\\ 10\\ 26\\ 26\\ 30\\ 24\\ 25
\end{pmatrix}
$$

18 equatioins
and
16 unknowns

In our case...

$$
\begin{pmatrix}
1111000000000000\\
0000111100000000\\
0000000011110000\\
0000000000001111\\
1000100010001000\\
0100010001000100\\
0010001000100010\\
0001000100010001\\
1000010000100001\\
0001001001001000\\
0100100000000000\\
0010010010000000\\
0010000100000000\\
0100001000010000\\
0000000000010010\\
0000000100100100\\
0000000010000100\\
0000100001000010
\end{pmatrix}
\begin{pmatrix}
x_1\\ x_2\\ x_3\\ x_4\\ x_5\\ x_6\\ x_7\\ x_8\\ x_9\\ x_{10}\\ x_{11}\\ x_{12}\\ x_{13}\\ x_{14}\\ x_{15}\\ x_{16}
\end{pmatrix}
=
\begin{pmatrix}
34\\ 34\\ 34\\ 34\\ 34\\ 34\\ 34\\ 34\\ 34\\ 34\\ 8\\ 21\\ 10\\ 26\\ 26\\ 30\\ 24\\ 25
\end{pmatrix}
$$

$$Ax = b$$

$$\arg\min_x \|Ax - b\|^2$$

Without regularisation

- Matriz mágica extraña

16	15	-10	13
-7	10	11	20
21	6	7	0
4	3	26	1

$$\arg\min_x \|Ax - b\|^2 + \lambda\|x\|^2$$

With regularisation

- Matriz mágica de Durero

16	3	2	13
5	10	11	8
9	6	7	12
4	**15**	**14**	1

The ubiquitous least squares solution...

$$Ax = b$$

Linear system solution

$$A^t A x = A^t b$$

$$x = \underbrace{(AA^t)^{-1} A^t}_{A^\dagger} b$$

$A^\dagger$ Pseudo-inverse of Penrose

property...

$$\arg\min_x \|Ax - b\|^2$$

Least-square solution

Property proof: version without/with regularisation

parámetro de regularización (no negativo)

$$\arg\min_x \underbrace{\|Ax - b\|^2 + \lambda\|Lx\|^2}_{J(x)}$$

$$\frac{dJ}{dx} = 0$$

$$(Ax - b, A\delta x) + \lambda(Lx, L\delta x) = 0$$

$$(A^t(Ax - b) + \lambda L^t Lx, \delta x) = 0$$

$$(A^t A + \lambda L^t L)x = A^t b$$

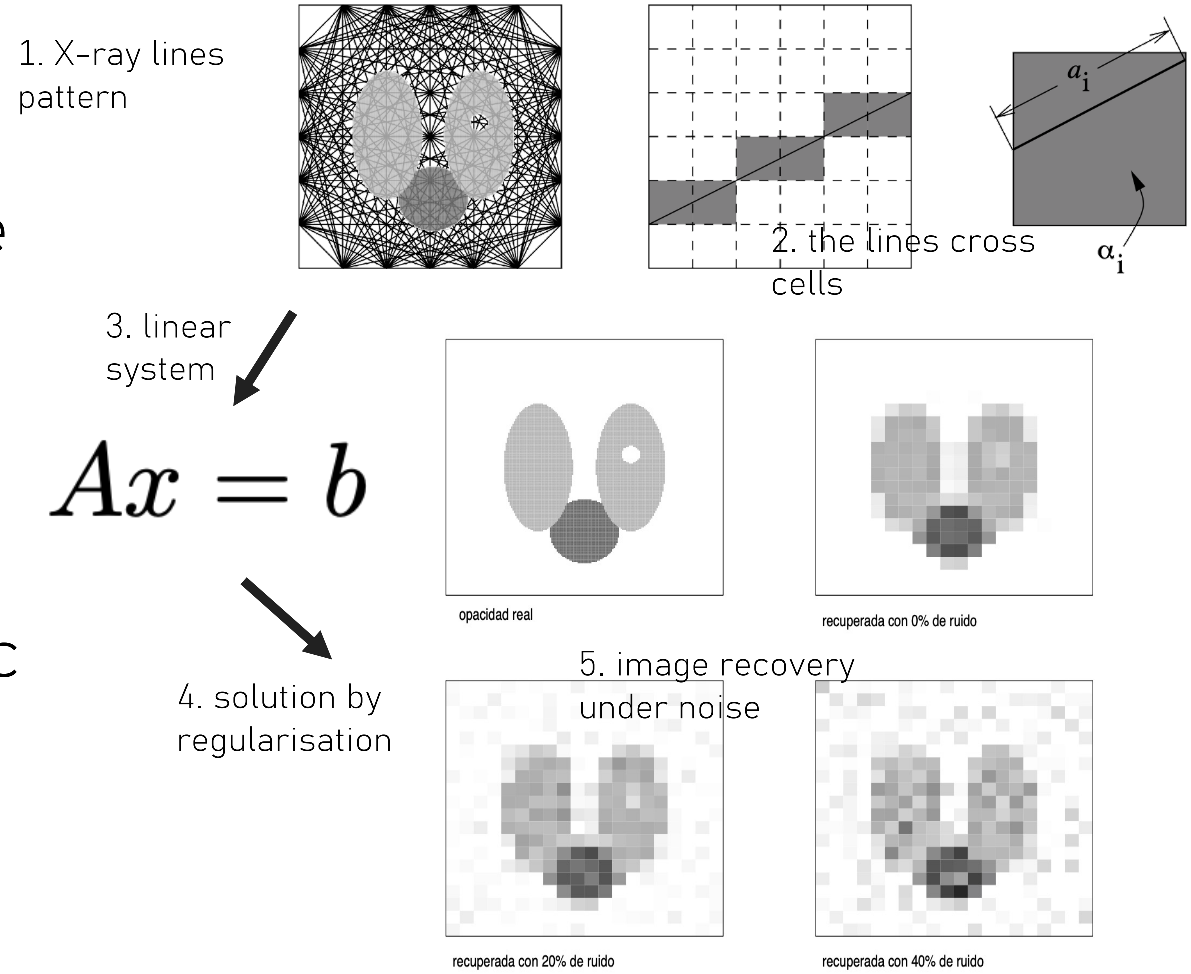
Without regularisation

$$x = \underbrace{(AA^t)^{-1} A^t}_{A^\dagger} b$$

With regularization

$$x = (AA^t + \lambda L^t L)^{-1} A^t b$$

Tomography: the equations are linear combinations of the attenuation/opacity values crossing the X-ray lines.



going beyond...

Bayes's theorem

$$\arg\max_x p(y|x) = \arg\max_x p(x|y)p(x)$$

$$= \arg\max_x \underbrace{\frac{1}{(2\pi)^{\frac{n}{2}}|R|} \exp\left(-\frac{1}{2}(Ax-y)^t R^{-1}(Ax-y)\right)}_{L(x)=p(y|x)}$$

model/data-consistency distribution

$$\underbrace{\frac{1}{(2\pi)^{\frac{n}{2}}|B|} \exp\left(-\frac{1}{2}(x-x_b)^t B^{-1}(x-x_b)\right)}_{\pi(x)=p(x)}$$

prior distribution

$$= \arg\min_x \underbrace{\underbrace{\frac{1}{2}(Ax-y)^t R^{-1}(Ax-y)}_{(1)} + \frac{1}{2}(x-x_b)^t B^{-1}(x-x_b)}_{(2)}$$

(1) ML estimator, least squares solution, data consistency

(2) MAP estimator, BLUE (Best Linear Unbiased Estimator) data assimilation, Kalman analysis, Minimum Variance, Regularised Least Squares

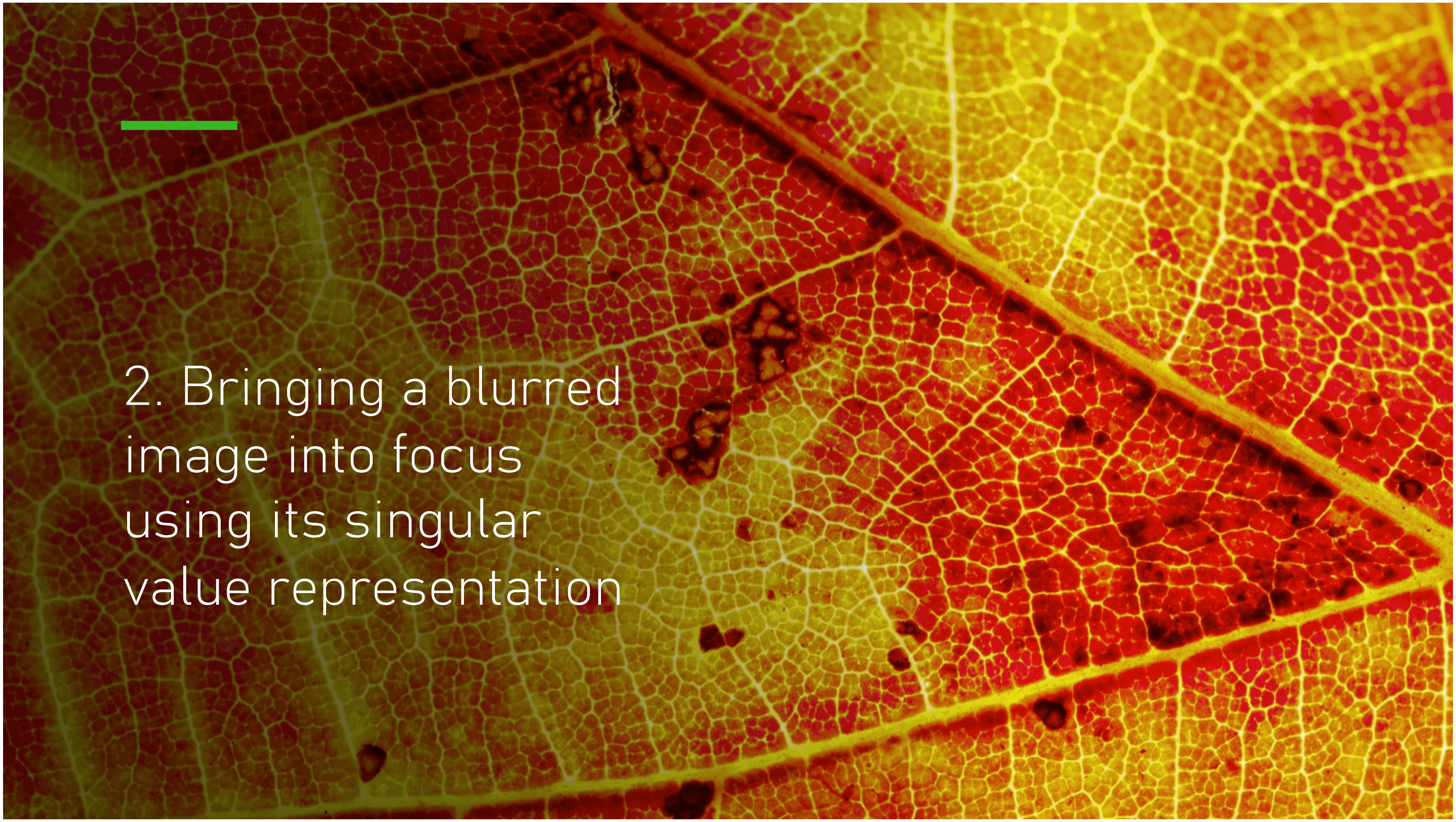
Regularisation parameter (scalar/diagonal case):

$$\lambda = \frac{\sigma_R^2}{\sigma_b^2}$$

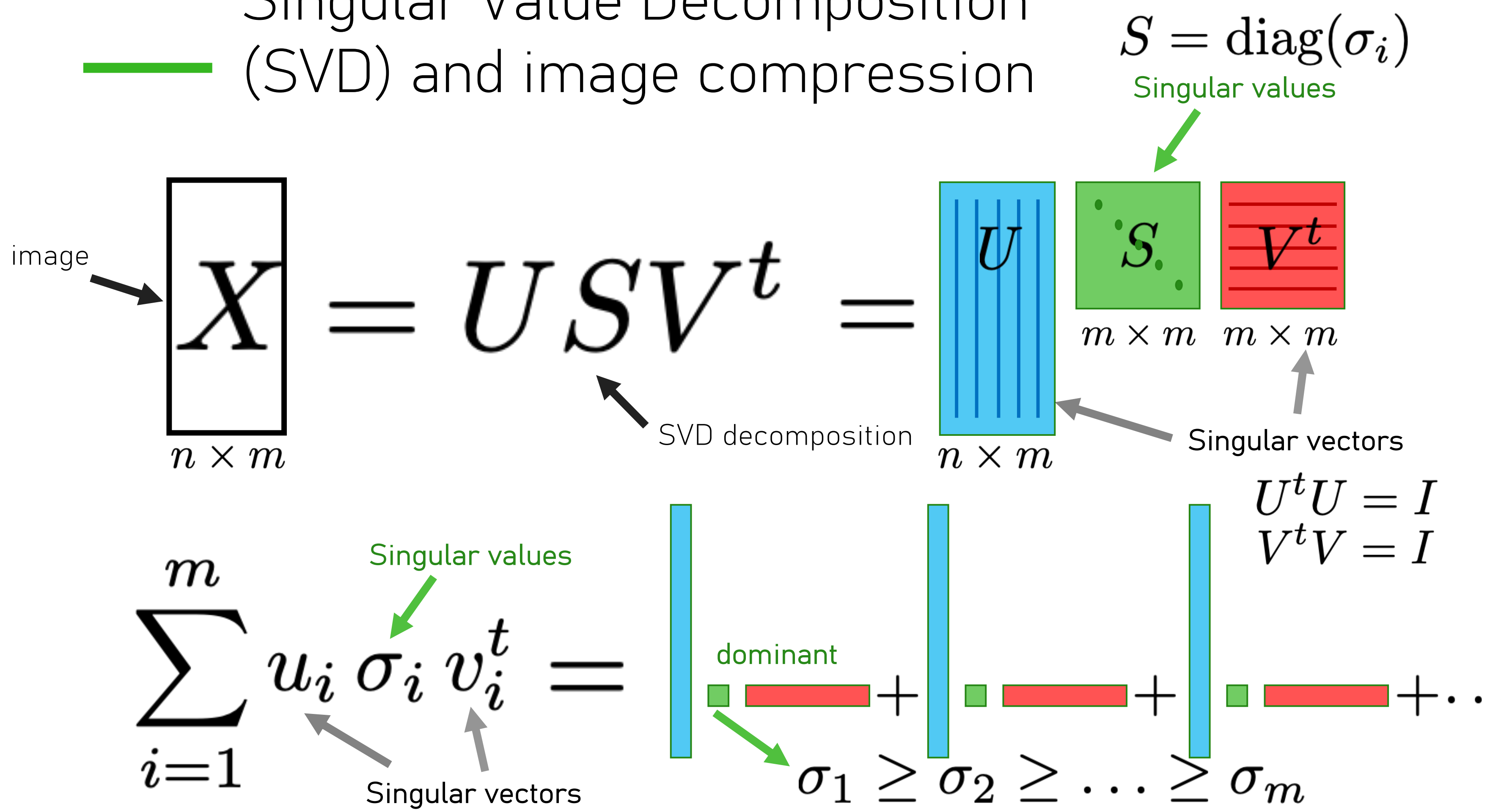
Analysis variance:

$$P^{-1} = B^{-1} + A^t R^{-1} A$$

2. Bringing a blurred image into focus using its singular value representation



Singular Value Decomposition (SVD) and image compression



Singular Value Decomposition (SVD)

```
A=imread('dog.png');
X=double(rgb2gray(A)); % Convert RBG->gray, 256 bit->double.
n = size(X,1);
m = size(X,2);
figure('Color','white')
subplot(1,4,1)
imagesc(X), axis off, axis equal, colormap gray
title(sprintf('original\n size %dx%d\n dim = %d\n 100%%',n, m, n*m),'FontSize',20);
[U,S,V] = svd(X,'econ');
i=1;
for r=[5 20 100]; % Truncation value
    i=i+1;
    Xapprox = U(:,1:r)*S(1:r,1:r)*V(:,1:r)';
    subplot(1,4,i)
    imagesc(Xapprox), axis off, axis equal, colormap gray
    title(sprintf('original\n rank=%d\n dim = %d\n %2.0f%%',r,(2*n+1)*r,100*(2*n+1)*r/n/m),'FontSize',20);
end
```

$$X = USV^t$$

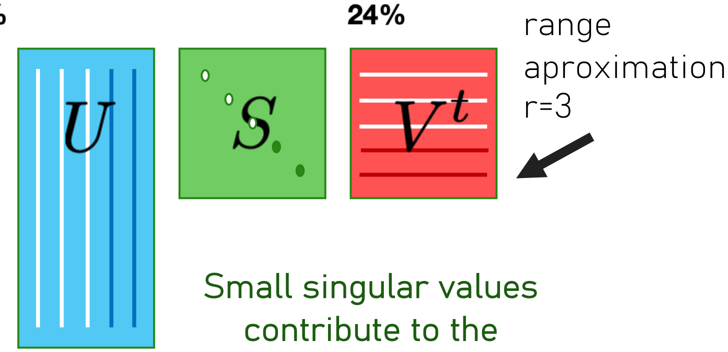
original
size 1152x818
dim = 942336
100%

original
rank=5
dim = 11525
1%

original
rank=20
dim = 46100
5%

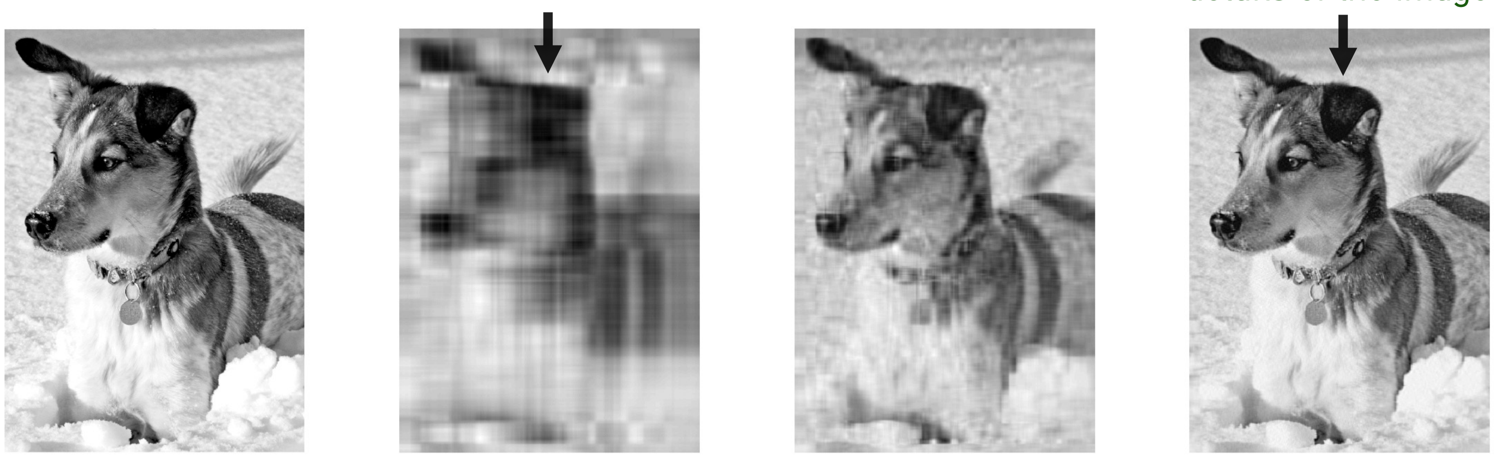
original
rank=100
dim = 230500
24%

$$X = USV^t$$

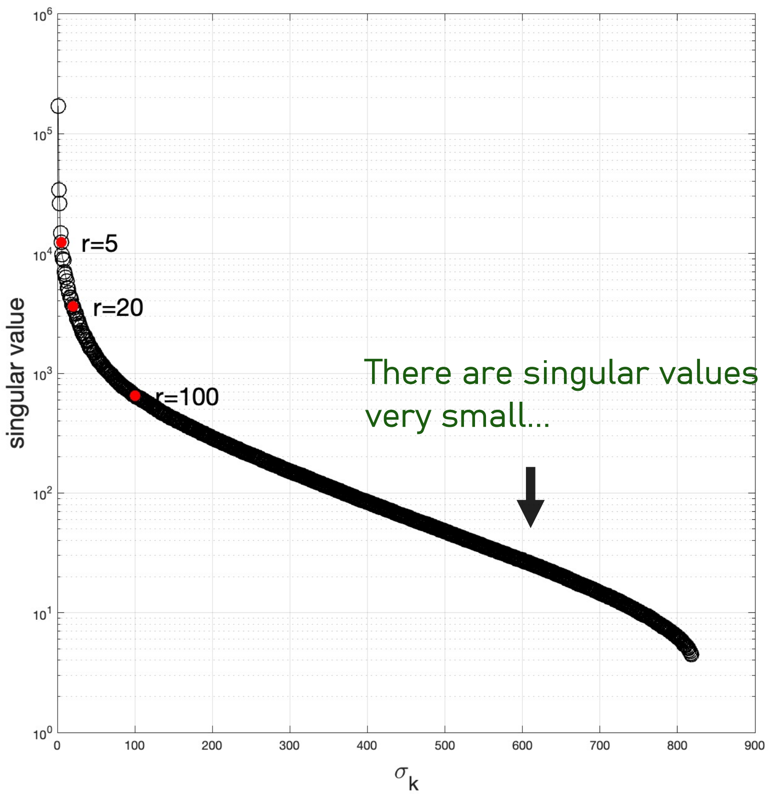


Large singular valuescontribute
to the structure of the image

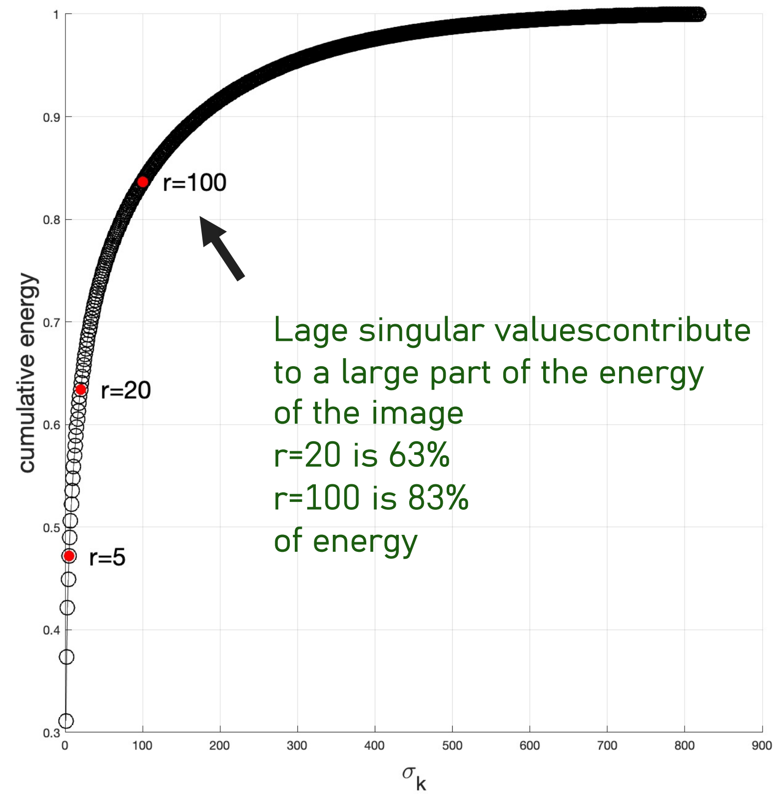
Small singular values
contribute to the
details of the image



Singular values ordered
from highest to lowest on a logarithmic scale



La energía acumulada por los 5, 20 y 10
primeros valores singulares



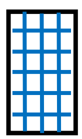
Example adapted from: Steven L. Brunton, J. Nathan Kutz. *Data Driven Science & Engineering. Machine Learning, Dynamical Systems and Control*, 2017

Singular value decomposition (SVD) for linear system

vector $nm \times 1$

Note: you must vectorise the image X as a vector x, the same for the data image B as a vector b

"reshape"



matrix $n \times m$

Model (diffusion, convolution) → $$Ax = b$$ ← data (blurred image)

Solution (original image vectorized) → x

$$A(x + \delta x) = b + \delta b$$

error in solution → δx; error in data → δb

Both properties indicate that small singular values (high frequencies) will lead to a poor approximation of the solution

SVD's properties:

$$A = USV^t$$

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\sigma_1}{\sigma_m} \frac{\|\delta b\|}{\|b\|}$$

conditioning and error

$$A^{-1} = VS^{-1}U^t = \sum_{i=1}^{m} v_i \frac{1}{\sigma_i} u_i^t$$

A inverse (if it were invertible)

Solution: truncated SVD (TSVD) and Tikhonov

Two alternatives for filtering high frequencies
high frequencies

Truncated sum

$$x \approx \widehat{A}^{\dagger} b = \sum_{i=1}^{r} v_i \frac{1}{\sigma_i} u_i^t b, \qquad r << m$$

range of the truncation

Approximate solution

Truncated pseudo-inverse

$$x \approx \sum_{i=1}^{m} \phi_i v_i \frac{1}{\sigma_i} u_i^t b \qquad \phi_i = \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2}$$

factor or filter

attenuates high frequencies

regularisation parameter

Explanation:SVD (remember that U and V are unitary matrices, their inverses are their transverses)

$$
\begin{aligned}
A^{\dagger} &= (A^t A)^{-1} A^t \\
&= (VSU^t USV^t)^{-1} VSU^t \\
&= VS^{-2} V^t VSU^t \\
&= VS^{-1} U^t \\
&= \sum_{i=1}^{m} v_i \frac{1}{\sigma_i} u_i^t \\
&= A^{-1}
\end{aligned}
$$

Valid even if A is not
is not invertible, not even
square! AtA must be
invertible otherwise it must b
regularise

Válida
solo si A es
cuadrada e invertible

"Rule of thumb": in reality, to solve a linear system in practice
NEVER invert the matrix of the system

References

- Steven L. Brunton, J. Nathan Kutz. *Data Driven Science & Engineering. Machine Learning, Dynamical Systems and Control*, 2017. Available online here
- Per Christian Hansen, *Discrete Inverse Problems: Insights and Algorithms*, SIAM, Philadelphia, 2010
- Mario Bertero, Patrizia Boccaci, Christine De Mol, *Introduction to Inverse Problems in Imaging*, Second Ed., CRC Press, Boca Raton, London, New York, 2021
- Andreas Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, Third Ed., Applied Mathematical Sciences Vol 120, Springer Nature Switzerland, 2021

thanks for your attention!
now the practice....