

Conditioning

GeoComput & ML

28 Apr. 2022

Transition



Transition



Transition

- Independent Thinking
- Innovative Solutions

Disruption

Logistics

who am I

- fellowship
- perspectives
- self-enrichment

Logistics

Interactions

- voice yourself
- in-class hours
- Matera times

Logistics

Class structure

- begins at each sharp hour
- duration : 45~50 min
- two breaks

Logistics

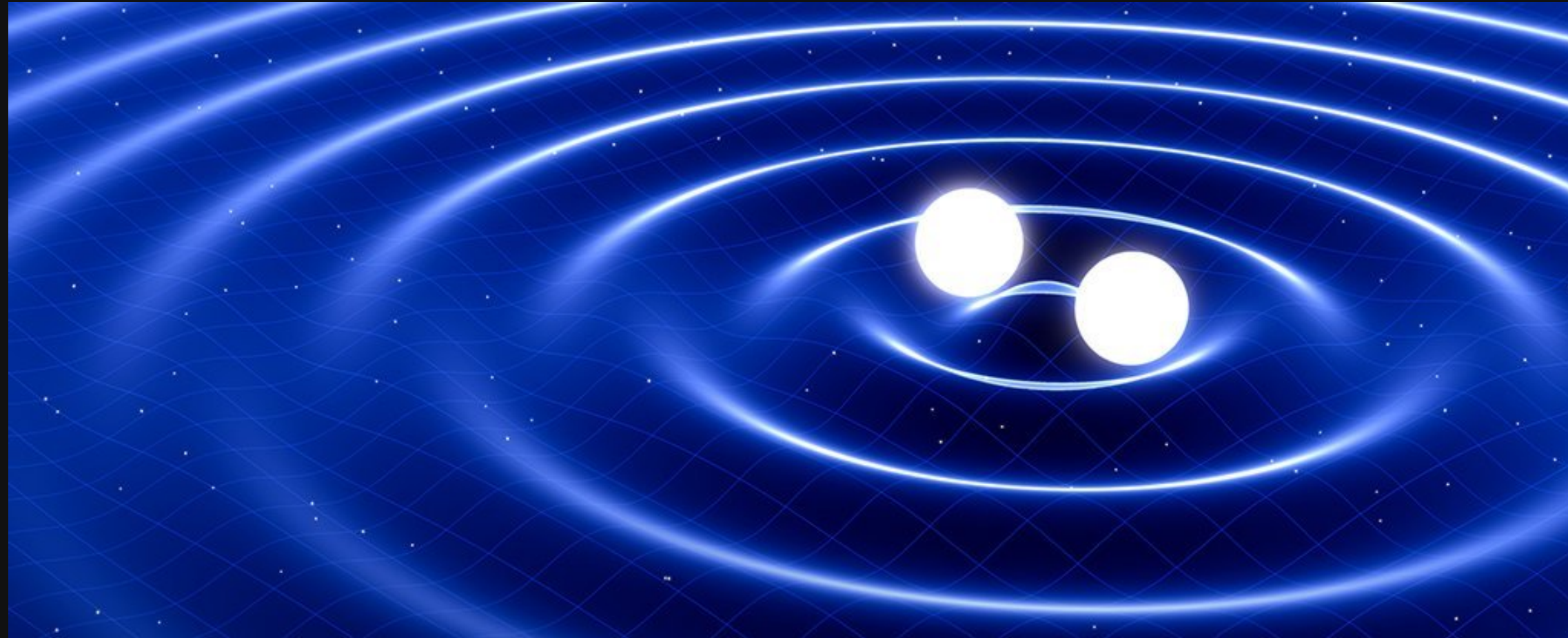
Questions

- in lecture notes
- respond to the interesting
- in class activity
- l.shen@spatial-ecology.net

Modelling

Motivation

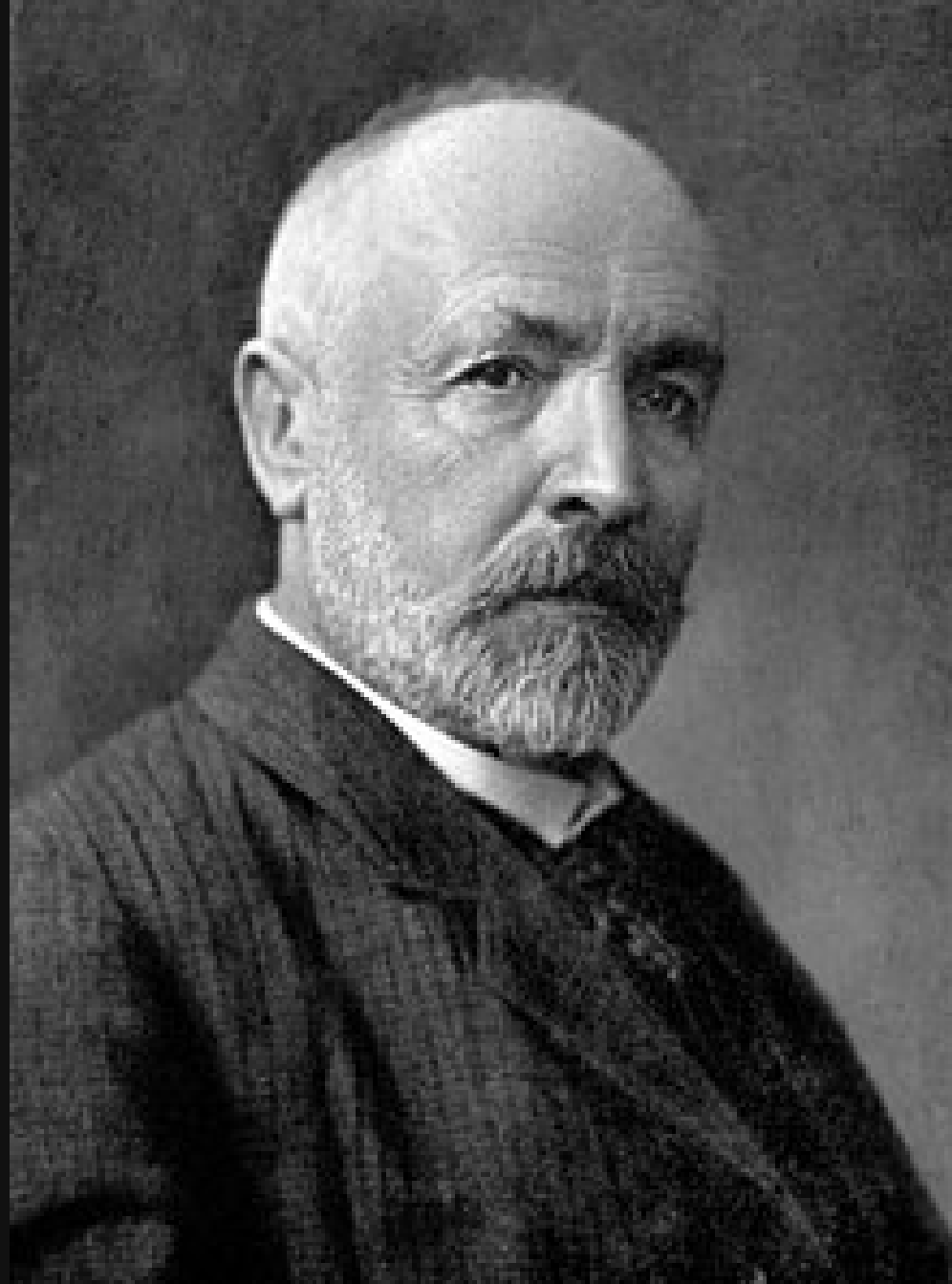
- do the impossible



Pillars

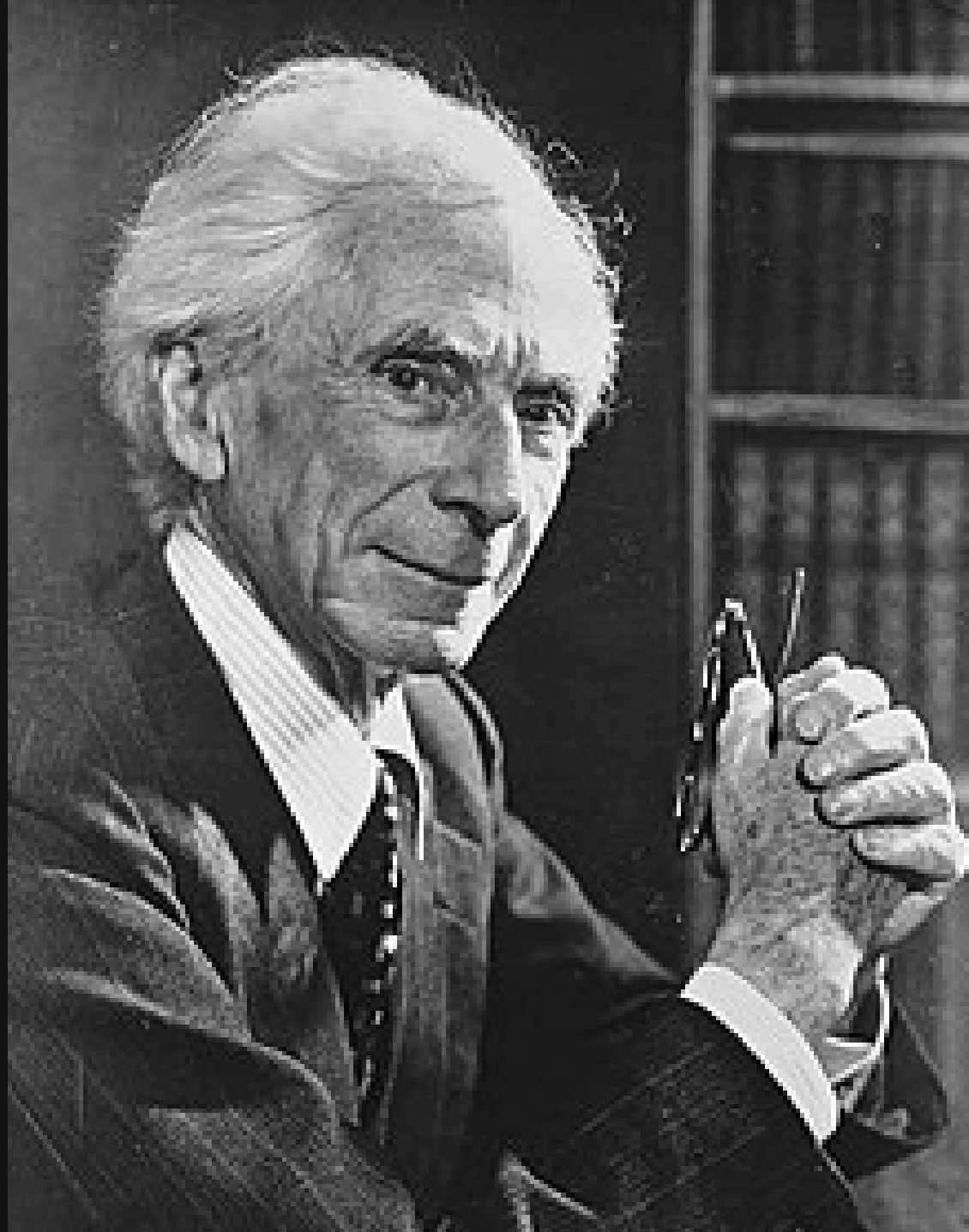
- Domain Knowledge
- Mathematics
- Computing

Set Theory



Georg Cantor

Set Theory



Bertrand Russell

Set Theory



A barber shaves those who only do NOT shave themselves

Incompleteness Theorem



Kurt Gödel

Computing

- Arithmetic
- Algorithms
- Analytics

Arithmetic

Definition

$$x = \pm \left(d_0 + \frac{d_1}{\beta^1} + \frac{d_2}{\beta^2} + \dots + \frac{d_{p-1}}{\beta^{p-1}} \right) \beta^E$$

β : base

$$0 \leq d_i \leq \beta - 1$$

p : precision

$$i = 0, \dots, p - 1$$

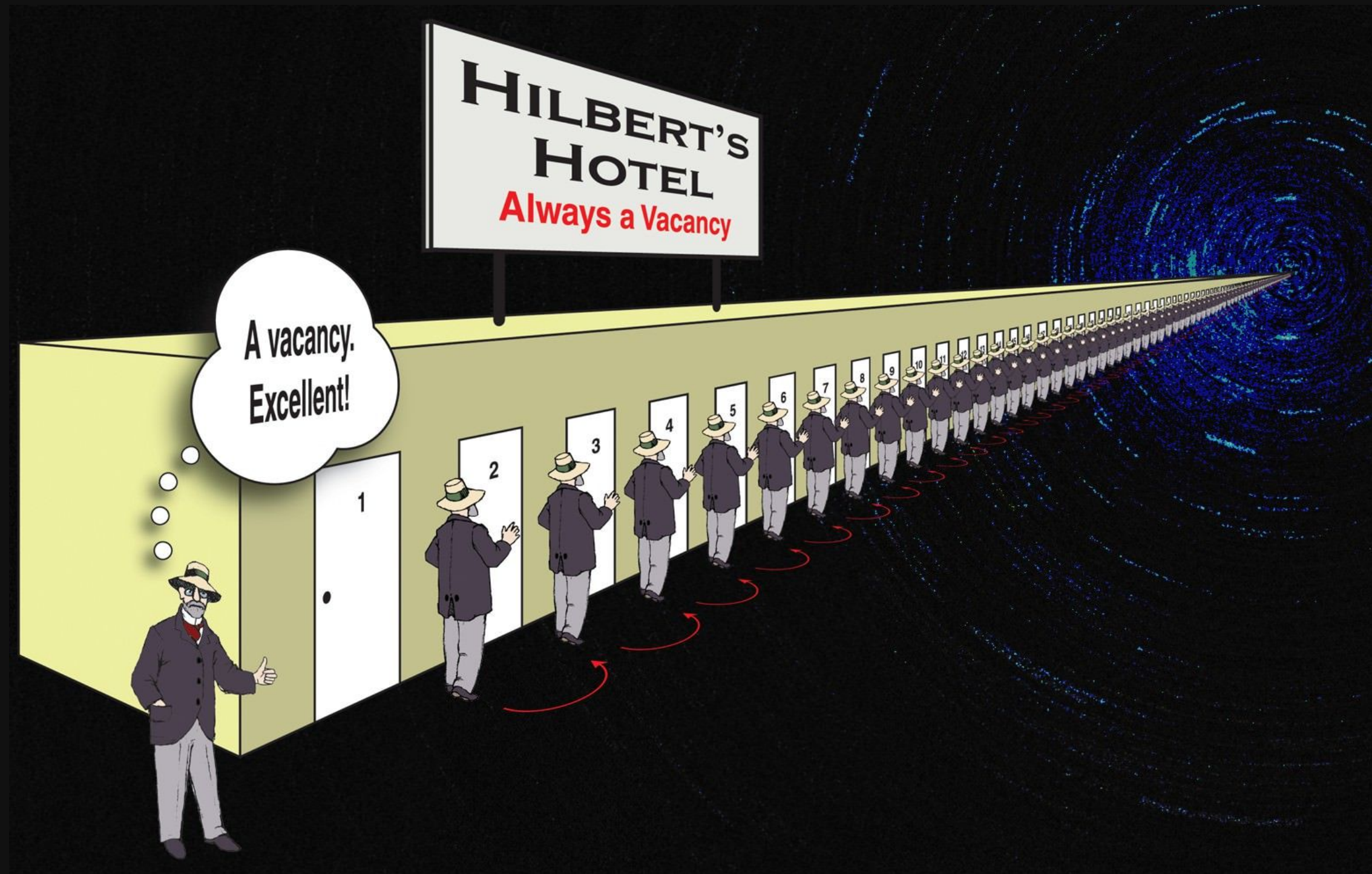
$[L, U]$: exponent range

$$E \in [L, U]$$

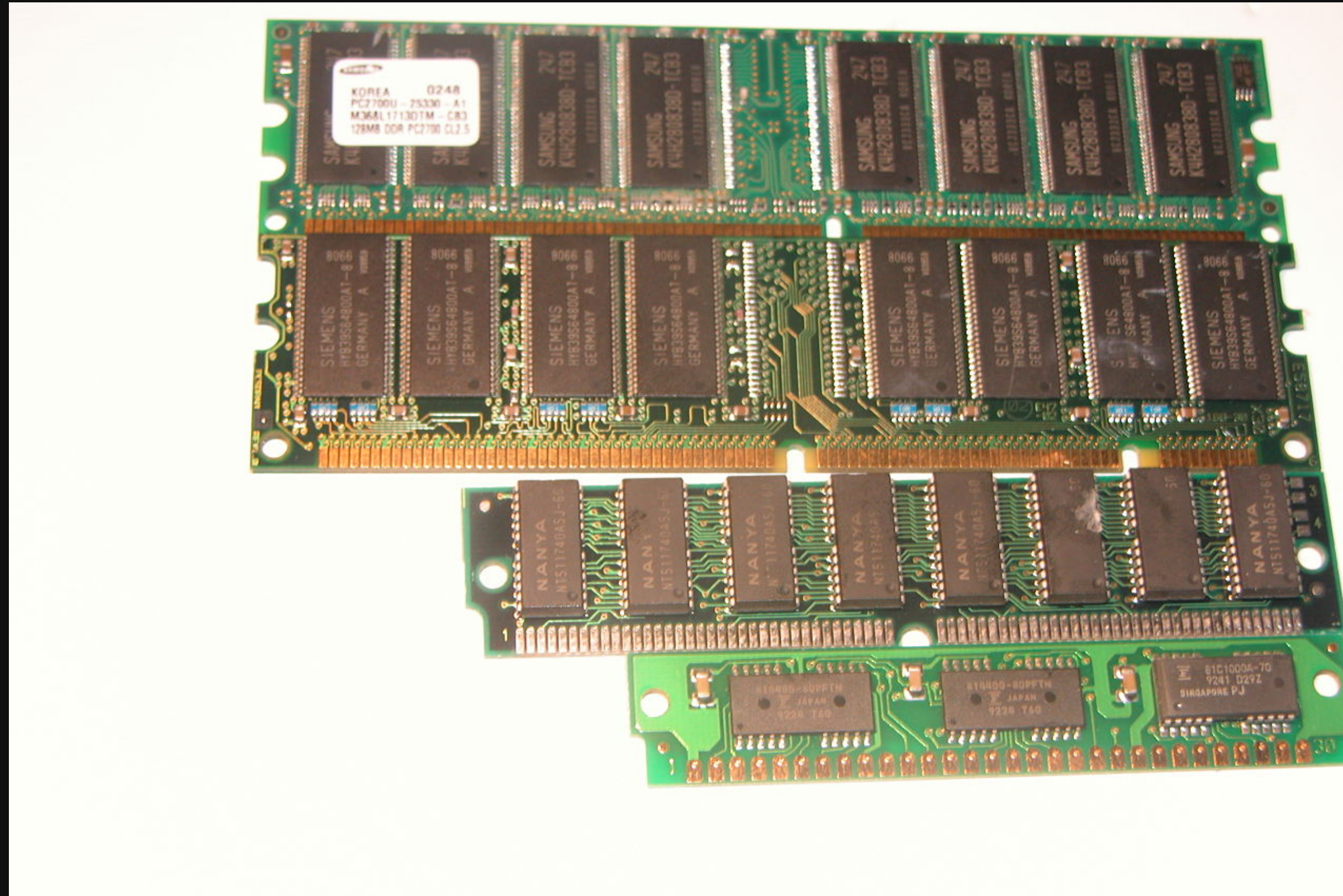
Machine Precision

finite and discrete

Machine Precision



Machine Precision



Machine Precision

let $x = 1/n$, $n \in \mathbb{Z}$, show $(n + 1)x - 1 = x$

```
for n in range(1 , 11) :  
    x = 1/n  
    xin = x  
    for k in range (30) :  
        x = (n + 1)*x - 1  
    print(n,xin,x)
```

```
1 1.0 1.0  
2 0.5 0.5  
3 0.3333333333333333 -21.0  
4 0.25 0.25  
5 0.2 6545103.021815777  
6 0.16666666666666666 -476641800.7969146  
7 0.14285714285714285 -9817068105.0  
8 0.125 0.125  
9 0.11111111111111111 4934324553889.695  
10 0.1 140892568471739.25
```

Machine Precision

$$\left(d_0 + \frac{d_1}{\beta^1} + \frac{d_2}{\beta^2} + \dots + \frac{d_{p-1}}{\beta^{p-1}} \right) \beta^E$$

$$\in [\beta^E, \beta^{(E+1)}]$$

Machine Precision

relative error

$$\in \left[\frac{(\beta/2)\beta^{-p}\beta^E}{\beta^{(E+1)}}, \frac{(\beta/2)\beta^{-p}\beta^E}{\beta^E} \right]$$

$$\in [(1/2)\beta^{-p}, (\beta/2)\beta^{-p}]$$

Machine Precision

therefore

$$\epsilon_{mach} = \beta^{1-p} / 2$$

Operations

$$\mathbf{fl}(x \text{ op } y) = (x \text{ op } y)(1 + \delta)$$

where $|\delta| \leq \epsilon_{mach}$,

\mathbf{fl} denotes floating representation and

op denotes any elementary arithmetic operations,
 $+$, $-$, \times and $/$.

Operations

Example

$$\begin{aligned}\mathbf{fl}(x(y+z)) &= \mathbf{fl}((x + (y+z)(1+\delta_1))(1+\delta_2)) \\ &= x(y+z)(1+\delta_1+\delta_2+\delta_1\delta_2) \\ &\approx x(y+z)(1+\delta_1+\delta_2) \\ &\leq x(y+z)(2\epsilon_{mach})\end{aligned}$$

Operations

```
>>> 3.14+3.14e-5  
3.1400314000000003
```


Operations

Catastrophic Cancellation

```
import math
def funexp(x,order) :
    ex = 1 + x
    for i in range(2 , order + 1):
        ex += math.pow(x,i)/math.factorial(i)
    return (ex)
ex = funexp(-2,4)
print(ex)
print(math.pow(math.e, -2))
```

0.33333333333333333337

0.1353352832366127

Operations

Computing Residuals

Suppose we obtained the solution \hat{x} for a linear system $ax = b$.

We are to compute the residual $r = b - a\hat{x}$

$$\mathbf{fl}(a\hat{x}) = a\hat{x}(1 + \delta_1)$$

$$\begin{aligned}\mathbf{fl}(b - a\hat{x}) &= (b - a\hat{x})(1 + \delta_1)(1 + \delta_2) \\ &= (r - a\hat{x}\delta_1)(1 + \delta_2) \\ &= r(1 + \delta_2) - a\hat{x}\delta_1 - a\hat{x}\delta_1\delta_2 \\ &\approx r(1 + \delta_2) - b\delta_1\end{aligned}$$

Conditioning

Question

well-posed, if its solution

- exists
- unique
- depends continuously on the data

Question

algorithm : stable

solution : well-conditioned

Errors

$$\begin{aligned}\text{total errors} &= \hat{f}(\hat{x}) - f(x) \\ &= (\hat{f}(\hat{x}) - f(\hat{x})) + (f(\hat{x}) - f(x)) \\ &= \text{computation error} + \text{data error}\end{aligned}$$

Errors

forward error : $\Delta y = \hat{y} - y$

backward error : $\Delta x = \hat{x} - x$, where $f(\hat{x}) = \hat{y}$

Errors

Example :

$$\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} \dots$$

$$\hat{y} = \hat{f}(x) = 1 - \frac{x^2}{2}$$

for $x = 1$, we have

$$\left. \begin{array}{l} y = f(1) \approx 0.5403 \\ \hat{y} = \hat{f}(1) = 0.5 \end{array} \right\} \Rightarrow \Delta y = \hat{y} - y = -0.0403$$

$$\Delta x = \hat{x} - x = \arccos(\hat{y}) - x = 0.0472$$

Condition number

a measure on the effects on the solution incurred
by data perturbation

$$\left| \frac{\Delta y / y}{\Delta x / x} \right| \approx \left| \frac{x f'(x)}{f(x)} \right|$$

Condition number

Question :

what is the condition number for the inverse function?

$$g(y) = f^{-1}(y)$$

Linear System

$$Ax = b$$

$$\text{cond}(\mathbf{A}) = \|\mathbf{A}\| \|\mathbf{A}^{-1}\|$$

Linear System

induced matrix norm

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

Linear System

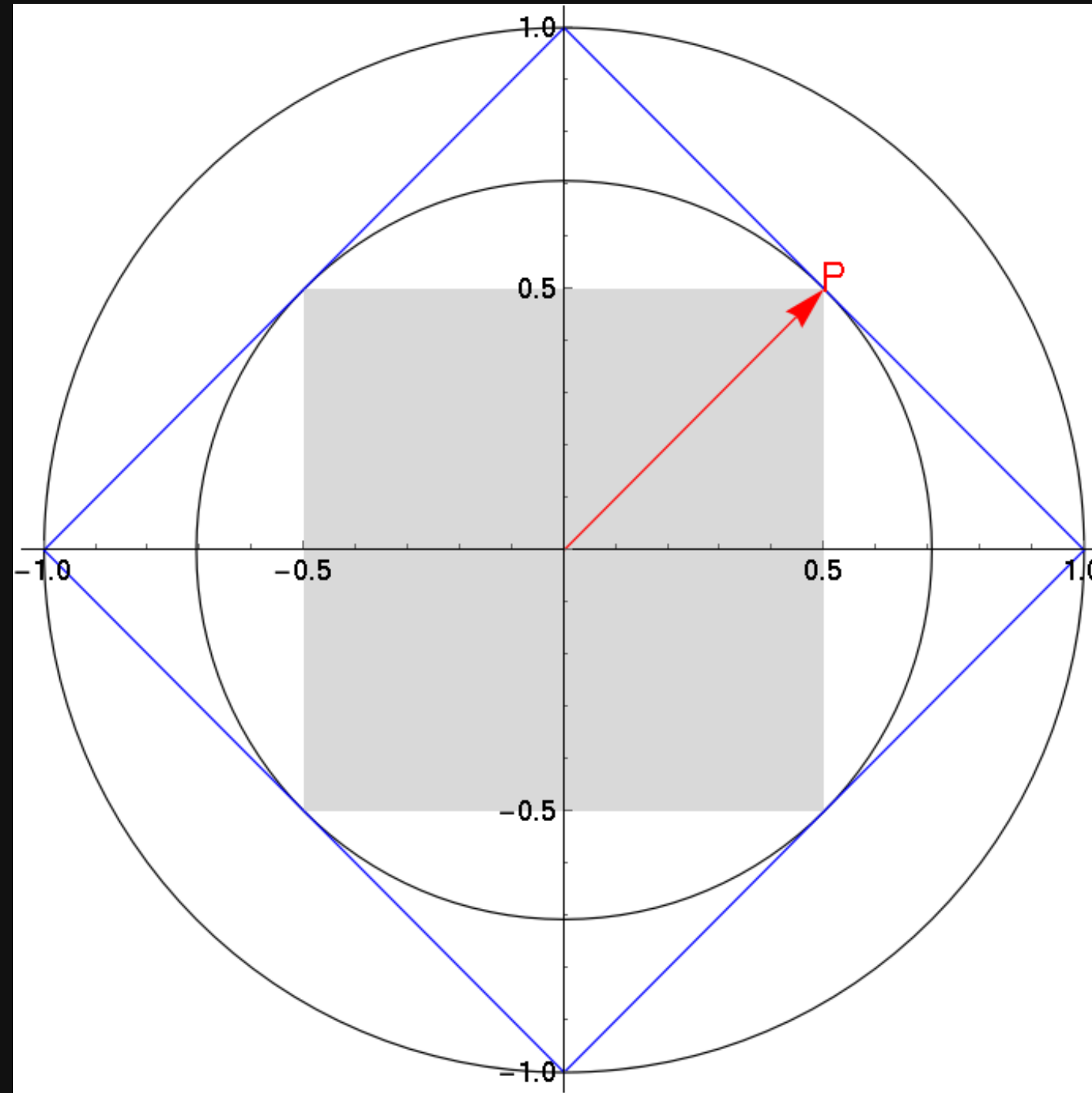
vector norms

$$\| \boldsymbol{x} \|_1 = \sum_i^n |x_i|$$

$$\| \boldsymbol{x} \|_2 = \left(\sum_i^n x_i^2 \right)^{1/2}$$

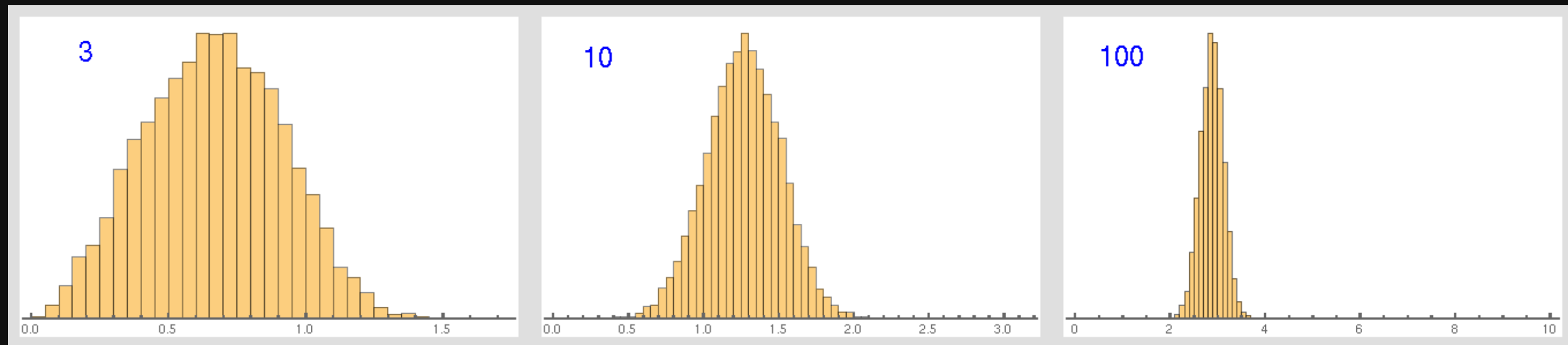
$$\| \boldsymbol{x} \|_\infty = \max_{1 \leq i \leq n} |x_i|$$

Linear System



Linear System

Euclidean distance between two random points



Linear System

linear system : $\mathbf{A}\mathbf{x} = \mathbf{b}$

residual : $\mathbf{r} = \mathbf{b} - \mathbf{A}\hat{\mathbf{x}}$

$$\begin{aligned}\|\Delta\mathbf{x}\| &= \|\hat{\mathbf{x}} - \mathbf{x}\| \\ &= \|\mathbf{A}^{-1}(\mathbf{A}\hat{\mathbf{x}} - \mathbf{b})\| \\ &= \|\mathbf{A}^{-1}\mathbf{r}\| \\ &\leq \|\mathbf{A}^{-1}\| \|\mathbf{r}\|\end{aligned}$$

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \text{cond}(\mathbf{A}) \frac{\|\mathbf{r}\|}{\|\mathbf{A}\| \|\hat{\mathbf{x}}\|}$$

Least Square

$$Ax \simeq b$$

Least Square

Normal equation

$$\begin{aligned}\phi(\boldsymbol{x}) &= (\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})^T (\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}) \\ &= \boldsymbol{b}^T \boldsymbol{b} - 2\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{b} + \boldsymbol{x}^T \boldsymbol{A}^T \boldsymbol{A}\boldsymbol{x}\end{aligned}$$

$$\mathbf{0} = \nabla \phi(\boldsymbol{x}) = 2\boldsymbol{A}^T \boldsymbol{A}\boldsymbol{x} - 2\boldsymbol{A}\boldsymbol{b}$$

$$\boldsymbol{A}^T \boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{b}$$

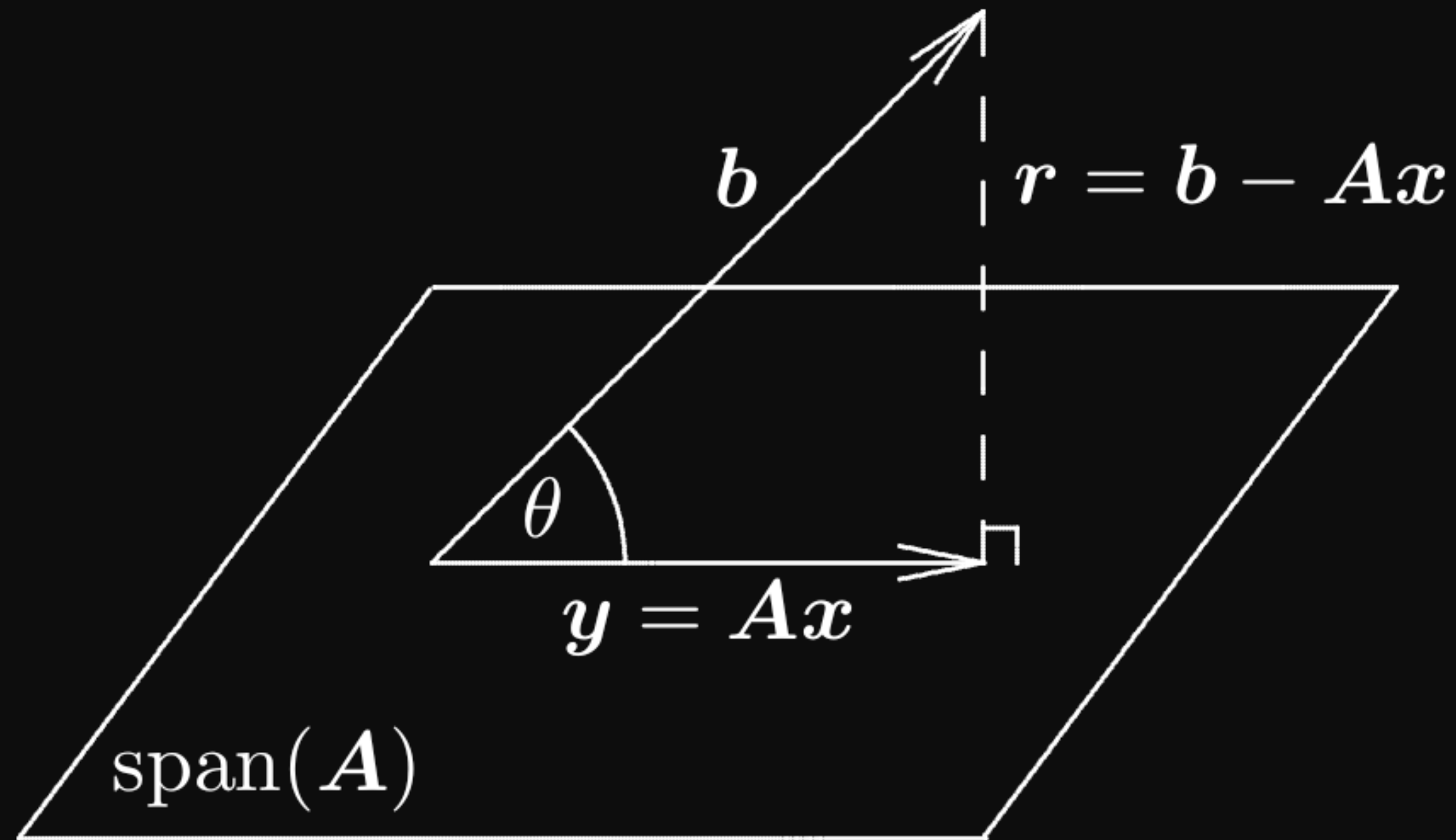
Least Square

Geometrical interpretation

$$\mathbf{y} = \mathbf{A}\mathbf{x} \in \text{span}(\mathbf{A})$$

orthogonal projection \mathbf{b} onto the $\text{span}(\mathbf{A})$

Least Square



Least Square

Projector matrix : idempotent

$$\mathbf{P}^2 = \mathbf{P}$$

Orthogonal projector :

$$\mathbf{P}^T = \mathbf{P}$$

$$\mathbf{P}_\perp = \mathbf{I} - \mathbf{P}$$

$$\mathbf{v} = (\mathbf{P} + (\mathbf{I} - \mathbf{P})\mathbf{v}) = \mathbf{P}\mathbf{v} + \mathbf{P}_\perp\mathbf{v}$$

Least Square

$$\begin{aligned}\| \mathbf{b} - \mathbf{Ax} \| &= \| \mathbf{P}(\mathbf{b} - \mathbf{Ax}) + \mathbf{P}_\perp(\mathbf{b} - \mathbf{Ax}) \|^2 \\ &= \| \mathbf{P}(\mathbf{b} - \mathbf{Ax}) \|^2 + \| \mathbf{P}_\perp(\mathbf{b} - \mathbf{Ax}) \|^2 \\ &= \| \mathbf{Pb} - \mathbf{Ax} \|^2 + \| \mathbf{P}_\perp \mathbf{b} \|^2\end{aligned}$$

$$\mathbf{Ax} = \mathbf{Pb}$$

Least Square

$$\mathbf{A}^T \mathbf{P} = \mathbf{A}^T \mathbf{P}^T = (\mathbf{P} \mathbf{A})^T = \mathbf{A}^T$$

$$\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{b}$$

$$\mathbf{P} = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$$

Least Square

Question :

Can you show \mathbf{P} is indeed a projection matrix?

Can \mathbf{P} be an identity matrix?

Least Square

pseudo inverse : $\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$

$$\text{cond}(\mathbf{A}) = \|\mathbf{A}\|_2 \|\mathbf{A}^+\|_2$$

Least Square

perturbation : $\mathbf{b} + \Delta \mathbf{b}$

$$\mathbf{A}^T \mathbf{A} \Delta \mathbf{x} = \mathbf{A}^T \Delta \mathbf{b}$$

$$\Delta \mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \Delta \mathbf{b} = \mathbf{A}^+ \Delta \mathbf{b}$$

$$\|\Delta \mathbf{x}\|_2 \leq \|\mathbf{A}^+\|_2 \|\Delta \mathbf{b}\|_2$$

$$\frac{\|\Delta \mathbf{x}\|_2}{\|\mathbf{x}\|_2} \leq \|\mathbf{A}^+\|_2 \frac{\|\Delta \mathbf{b}\|_2}{\|\mathbf{b}\|_2}$$

$$= \text{cond}(\mathbf{A}) \frac{\|\mathbf{b}\|_2}{\|\mathbf{Ax}\|_2} \frac{\|\Delta \mathbf{b}\|_2}{\|\mathbf{b}\|_2}$$

$$\leq \text{cond}(\mathbf{A}) \frac{\|\mathbf{b}\|_2}{\|\mathbf{Ax}\|_2} \frac{\|\Delta \mathbf{b}\|_2}{\|\mathbf{b}\|_2}$$

$$= \text{cond}(\mathbf{A}) \frac{1}{\cos(\theta)} \frac{\|\Delta \mathbf{b}\|_2}{\|\mathbf{b}\|_2}$$

Least Square

Example

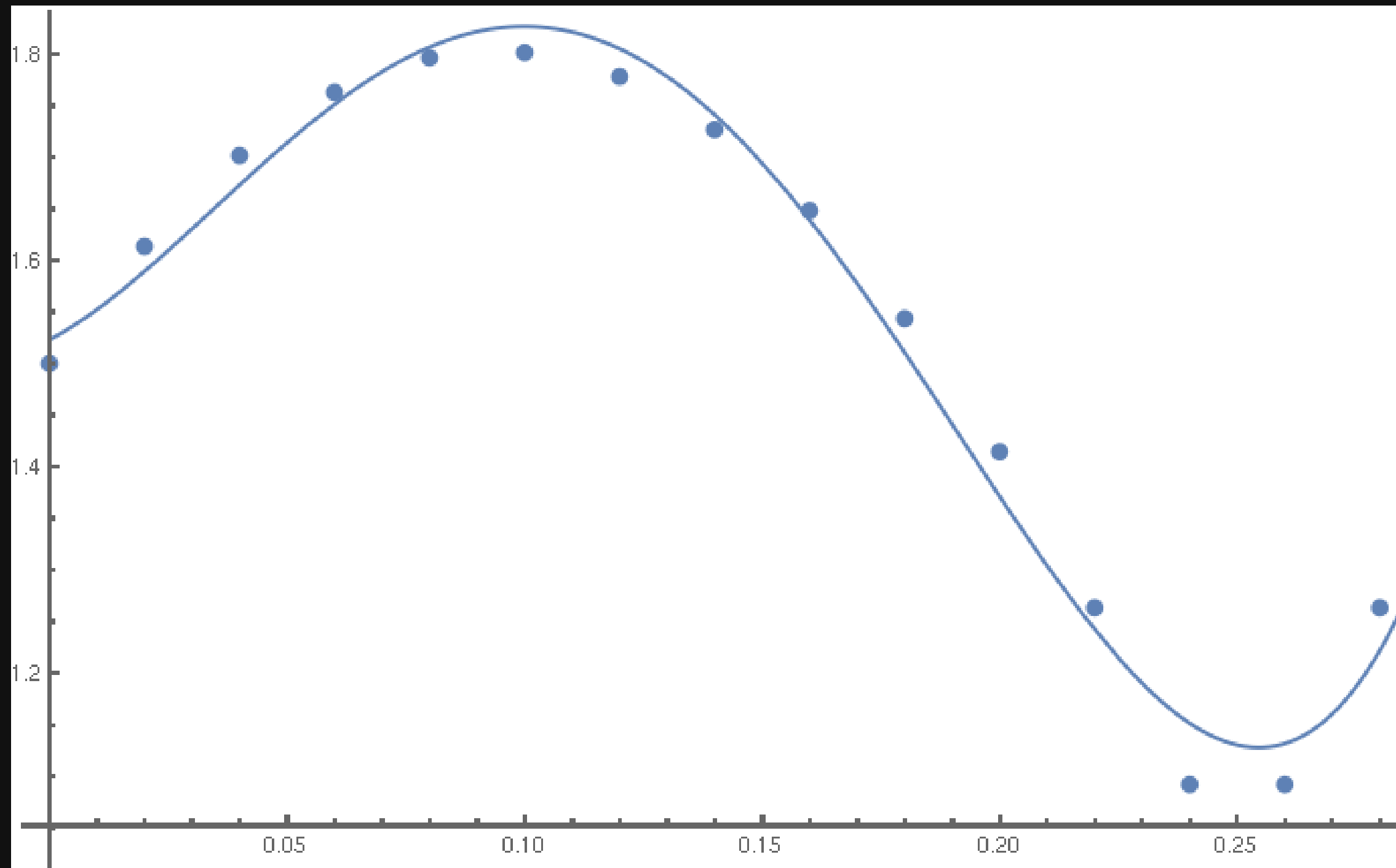
A 4th order polynomial fit

0.	1.5
0.02	1.61351
0.04	1.70156
0.06	1.76279
0.08	1.79621
0.1	1.80131
0.12	1.778
0.14	1.72665
0.16	1.64807
0.18	1.5435
0.2	1.41458
0.22	1.26336
0.24	1.09221
0.26	1.09221
0.28	1.26336

Least Square

1	0.	0.	0.	0.
1	0.02	0.0004	$8. \times 10^{-6}$	1.6×10^{-7}
1	0.04	0.0016	0.000064	2.56×10^{-6}
1	0.06	0.0036	0.000216	0.00001296
1	0.08	0.0064	0.000512	0.00004096
1	0.1	0.01	0.001	0.0001
1	0.12	0.0144	0.001728	0.00020736
1	0.14	0.0196	0.002744	0.00038416
1	0.16	0.0256	0.004096	0.00065536
1	0.18	0.0324	0.005832	0.00104976
1	0.2	0.04	0.008	0.0016
1	0.22	0.0484	0.010648	0.00234256
1	0.24	0.0576	0.013824	0.00331776
1	0.26	0.0676	0.017576	0.00456976
1	0.28	0.0784	0.021952	0.00614656

Least Square



Least Square

$$\text{cond}(\mathbf{A}) = 7.6 \times 10^5$$

$$\cos(\theta) = \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{b}\|_2} \approx 0.99$$

Take-home



Reading Assignment

What Role Does Hydrological Science Play in the Age
of Machine Learning?

Grey S. Nearing^{1,2}, Frederik Kratzert³, Alden Keefe Sampson¹, Craig S.
Pelissier⁴, Daniel Klotz³, Jonathan M. Frame², Cristina Prieto⁵, Hoshin V.
Gupta⁶

Acknowledgement

Thanks for Your Attention

References

- M. Holmes, Introduction to Scientific Computing and Data Analysis, 2016
- B. Gustafsson, Scientific computing from a historical perspective, 2010
- M. Heath, Scientific Computing An Introductory Survey, 2018
- Goldberg, David. ACM Computing Surveys. 1991, 23 (1): 5–48.
- A. Agresti, Foundations of Linear and Generalized Linear Models, 2015
- C. McCulloch, et.al., Generalised, Linear and Mixed Models, 2008
- M. Greenberg, Advanced Engineering Mathematics, 2004
- <https://www.sciencealert.com/gravitational-waves>
- <https://www.pinterest.com/pin/356347389250592062>
- https://en.wikipedia.org/wiki/Georg_Cantor
- https://en.wikipedia.org/wiki/Kurt_G%C3%B6del
- https://en.wikipedia.org/wiki/Bertrand_Russell
- <https://ieeexplore.ieee.org/document/4610935>