Lecture Two: What to Learn

Yi Ma

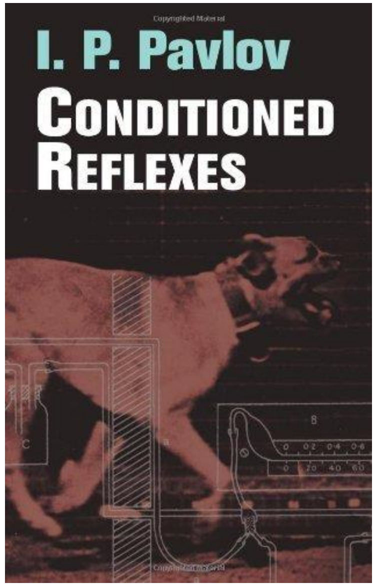
Director of the School of Computing and Data Science
Director of the Institute of Data Science

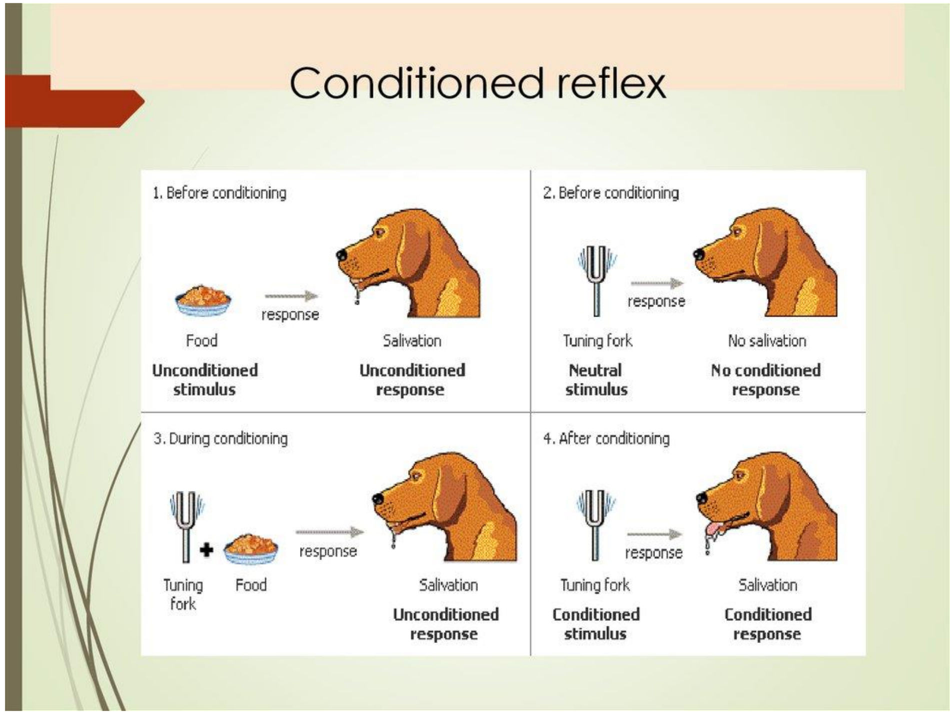
What to Learn?

Predictable information from data sensed of the external world

(Humans, including animals, have learned a model of the world)

Prediction based on correlation among things

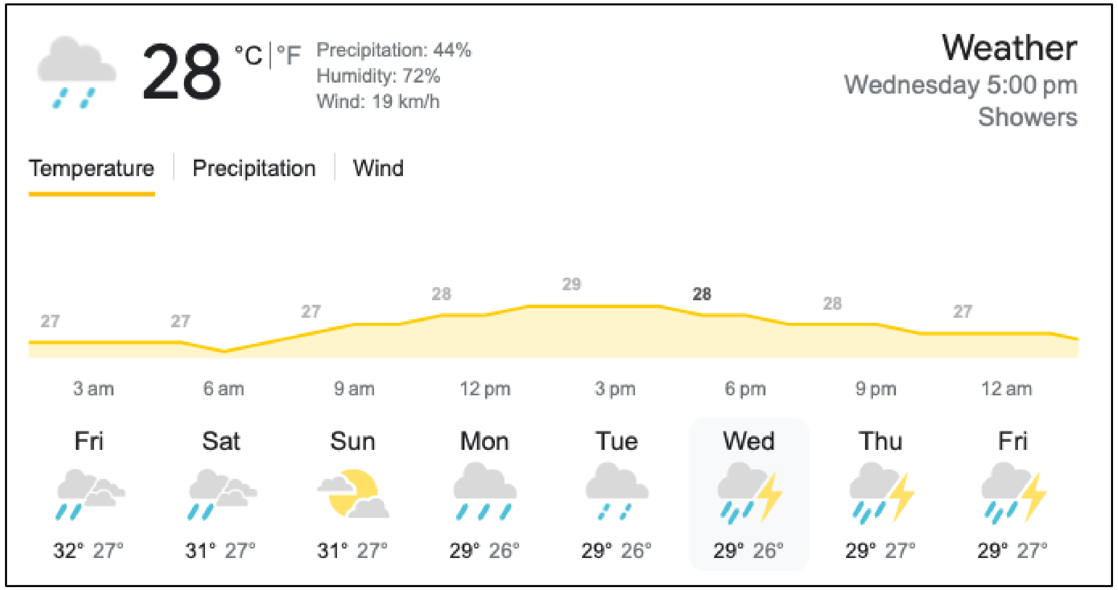




What to Learn?

Predictable information from data sensed of the external world

(Humans, including animals, have learned a model of the world)

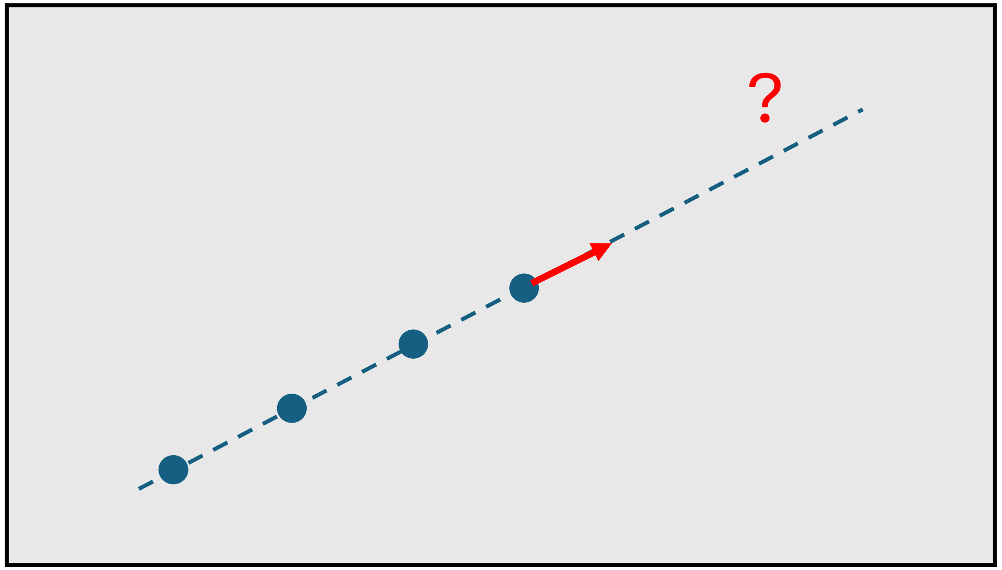


But correlation is not causality.

What to Learn? **Predictable information** from data sensed of the external world

(Human starts to develop physical and mathematical models of the world)

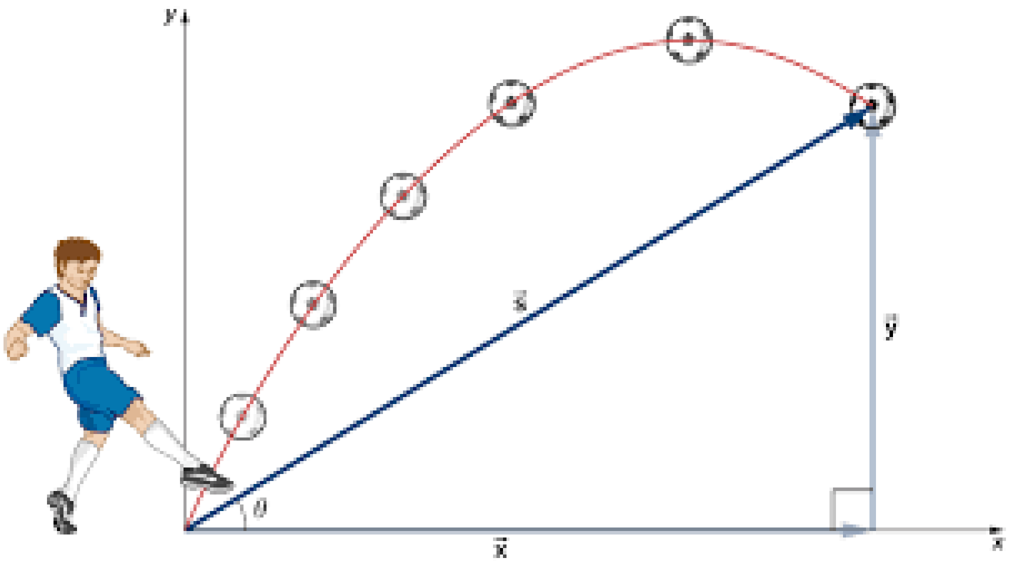
Newton’s first law of Motion

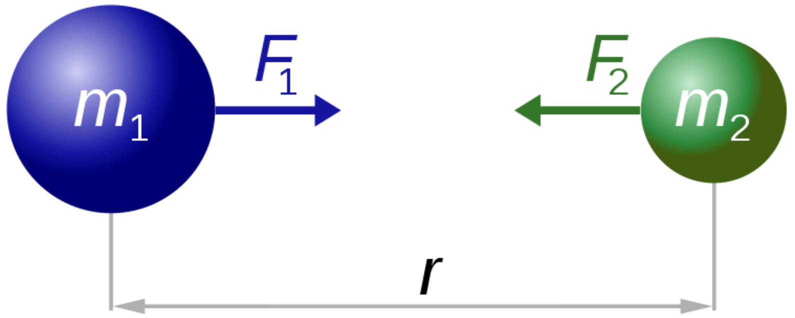


What to Learn? **Predictable information** from data sensed of the external world

(Human starts to develop physical and mathematical models of the world)

Newton's second law of motion and law of gravity





$$F_1 = F_2 = G\frac{m_1 \times m_2}{r^2}$$

What to Learn? **Predictable information** from data sensed of the external world

(Human starts to develop physical and mathematical models of the world)

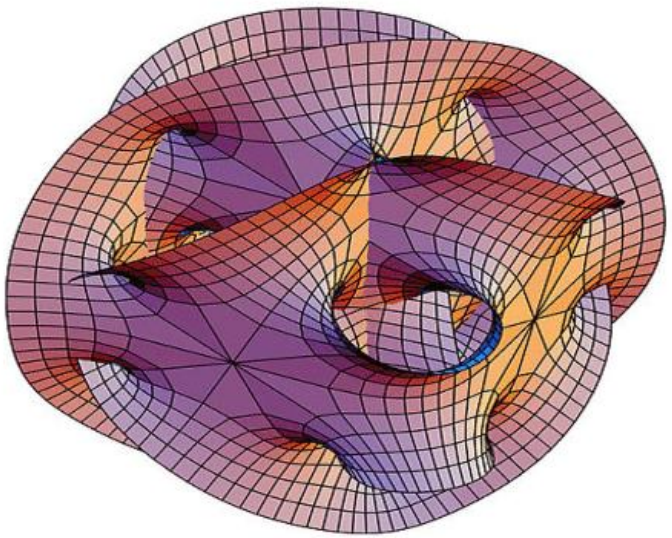
Why equations?

17 Equations That Changed the World
by Ian Stewart

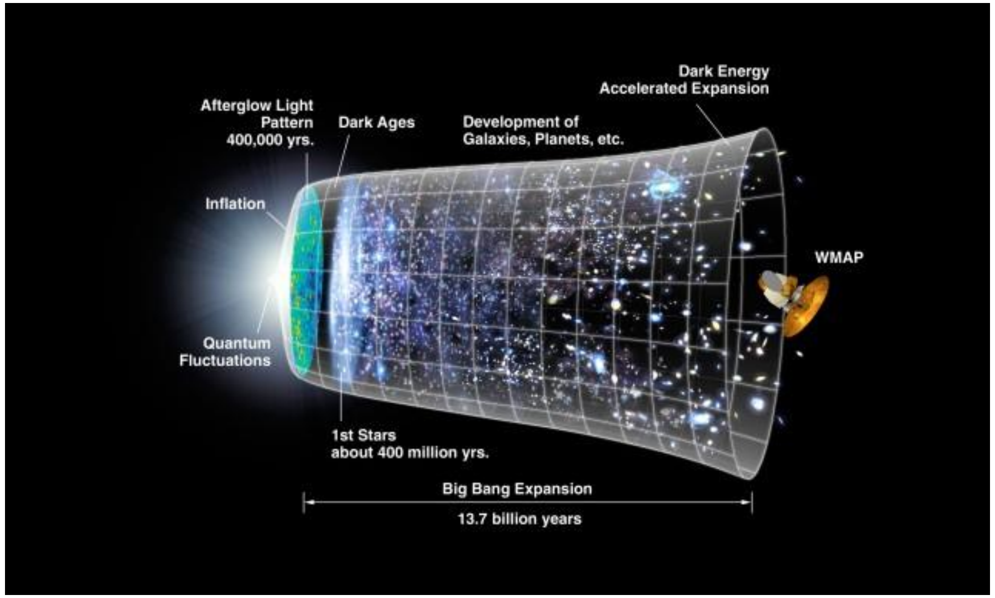
1.	**Pythagoras's Theorem**	$a^2 + b^2 = c^2$	Pythagoras,530 BC
2.	**Logarithms**	$\log xy = \log x + \log y$	John Napier, 1610
3.	**Calculus**	$\frac{\mathrm{d}f}{\mathrm{d}t} = \lim_{h \to 0} = \frac{f(t+h) - f(t)}{h}$	Newton, 1668
4.	**Law of Gravity**	$F = G\frac{m_1 m_2}{r^2}$	Newton, 1687
5.	**The Square Root of Minus One**	$i^2 = -1$	Euler, 1750
6.	**Euler's Formula for Polyhedra**	$V - E + F = 2$	Euler, 1751
7.	**Normal Distribution**	$\Phi(x) = \frac{1}{\sqrt{2\pi\rho}} e^{\frac{(x-\mu)^2}{2\rho^2}}$	C.F. Gauss, 1810
8.	**Wave Equation**	$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$	J. d'Almbert, 1746
9.	**Fourier Transform**	$f(\omega) = \int_{\infty}^{\infty} f(x) e^{-2\pi i x \omega} \mathrm{d}x$	J. Fourier, 1822
10.	**Navier-Stokes Equation**	$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v}\right) = -\nabla p + \nabla \cdot \mathbf{T} + \mathbf{f}$	C. Navier, G. Stokes, 1845
11.	**Maxwell's Equations**	$\nabla \cdot \mathbf{E} = 0$, $\nabla \cdot \mathbf{H} = 0$, $\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t}$, $\nabla \times \mathbf{H} = \frac{1}{c}\frac{\partial E}{\partial t}$	J.C. Maxwell, 1865
12.	**Second Law of Thermodynamics**	$\mathrm{d}S \geq 0$	L. Boltzmann, 1874
13.	**Relativity**	$E = mc^2$	Einstein, 1905
14.	**Schrodinger's Equation**	$i\hbar\frac{\partial}{\partial t}\Psi = H\Psi$	E. Schrodinger, 1927
15.	**Information Theory**	$H = -\sum p(x) \log p(x)$	C. Shannon, 1949
16.	**Chaos Theory**	$x_{t+1} = kx_t(1 - x_t)$	Robert May, 1975
17.	**Black-Scholes Equation**	$\frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} - rV = 0$	F. Black, M. Scholes, 1990

What to Learn? **Predictable information** from data sensed of the external world

The Calabi-Yau manifold
(the standard model & string theory)



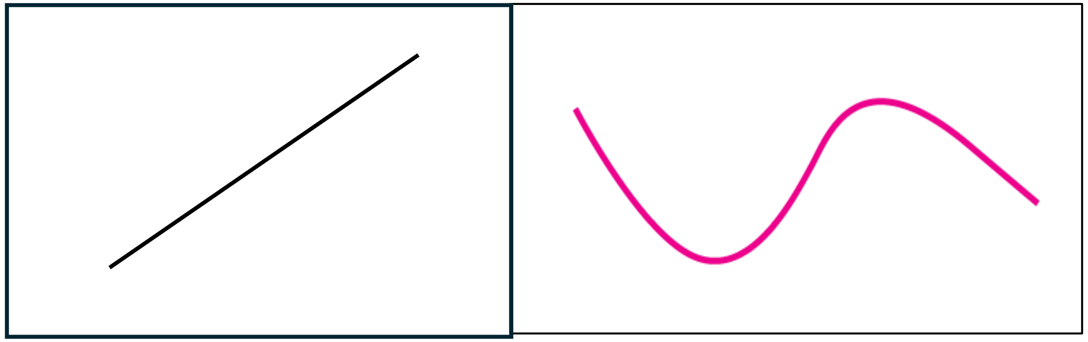
9+1-dimensional universe



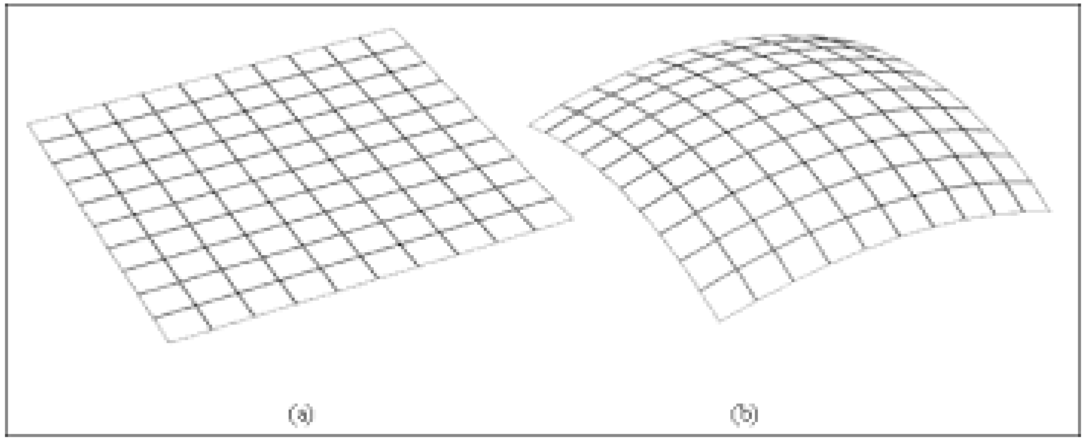
What to Learn?

How to model predictability mathematically or computationally?

$$x_{n+1} = f(x_n)$$



$$x_{n+1} = f(x_n, x_{n-1})$$



What to Learn?

How to model predictability mathematically or computationally?

$$x_{n+1} = f(x_n, x_{n-1}, \dots, x_{n-d+1})$$

What to Learn?

How to model predictability mathematically or computationally?

$$x_{n+1} = f(x_n, u_n)$$

What to Learn? How to model predictability mathematically or computationally?

$$x_{n+1} = f(x_n, x_{n-1}, \ldots, x_{n-d+1})$$

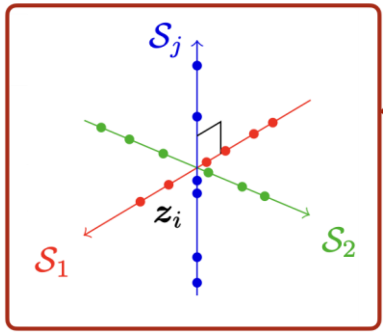
$$\vec{x} = [x_i, x_{i+1}, \ldots, x_{i+D-1}]$$

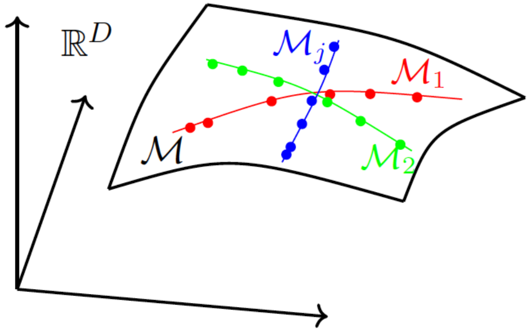
What to Learn? How to model predictability mathematically or computationally?

$$x_{n+1} = f_1(x_n, x_{n-1}, \ldots, x_{n-d+1})$$

$$x_{n+1} = f_2(x_n, x_{n-1}, \ldots, x_{n-d+1})$$

$$x_{n+1} = f_3(x_n, x_{n-1}, \ldots, x_{n-d+1})$$

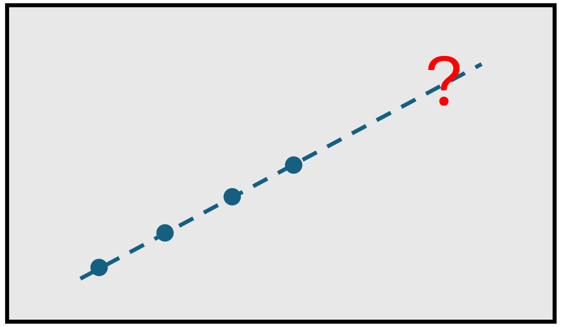


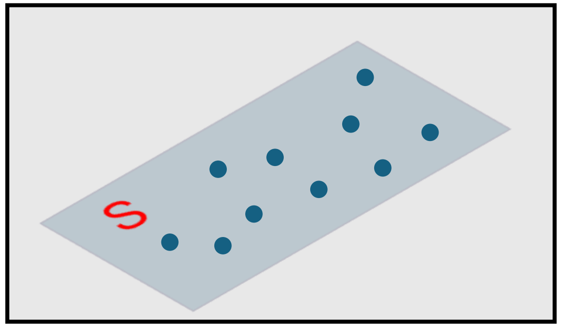


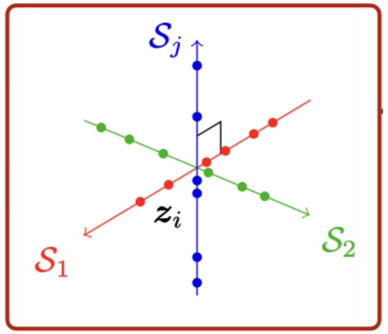
$$\vec{x} = [x_i, x_{i+1}, \ldots, x_{i+D-1}]$$

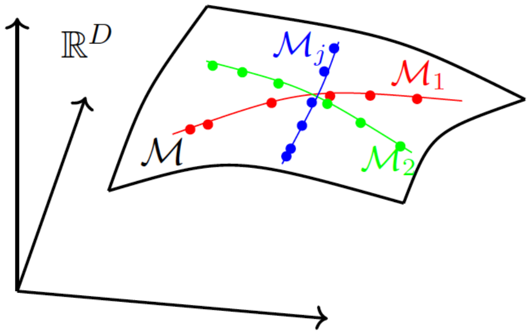
What to Learn? **Predictable information** from data sensed of the external world

Mathematically, all predictable information can be modeled as certain low-dimensional structures in the high-dimensional data



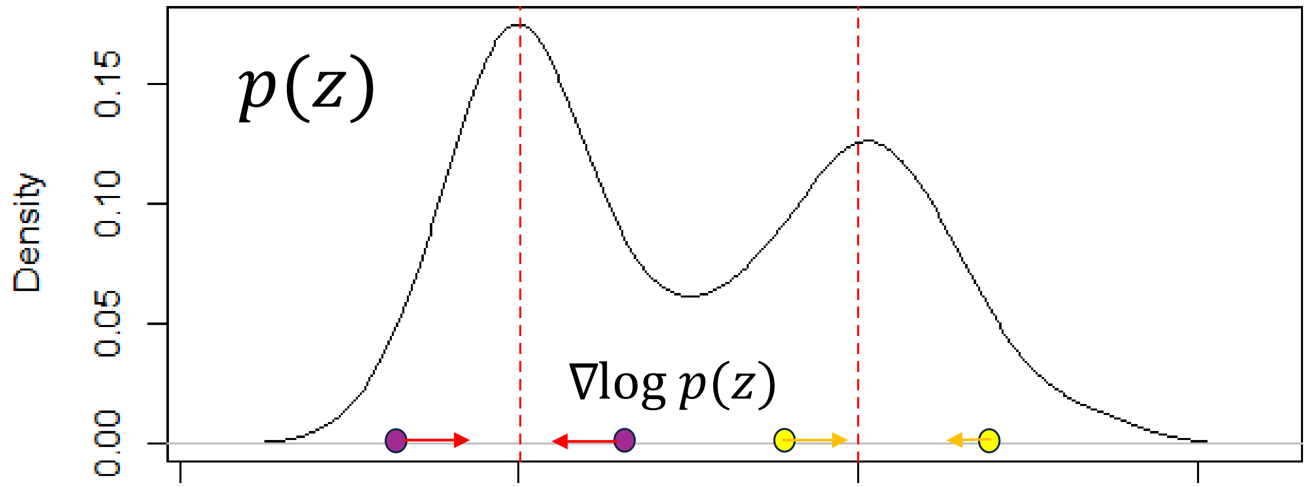


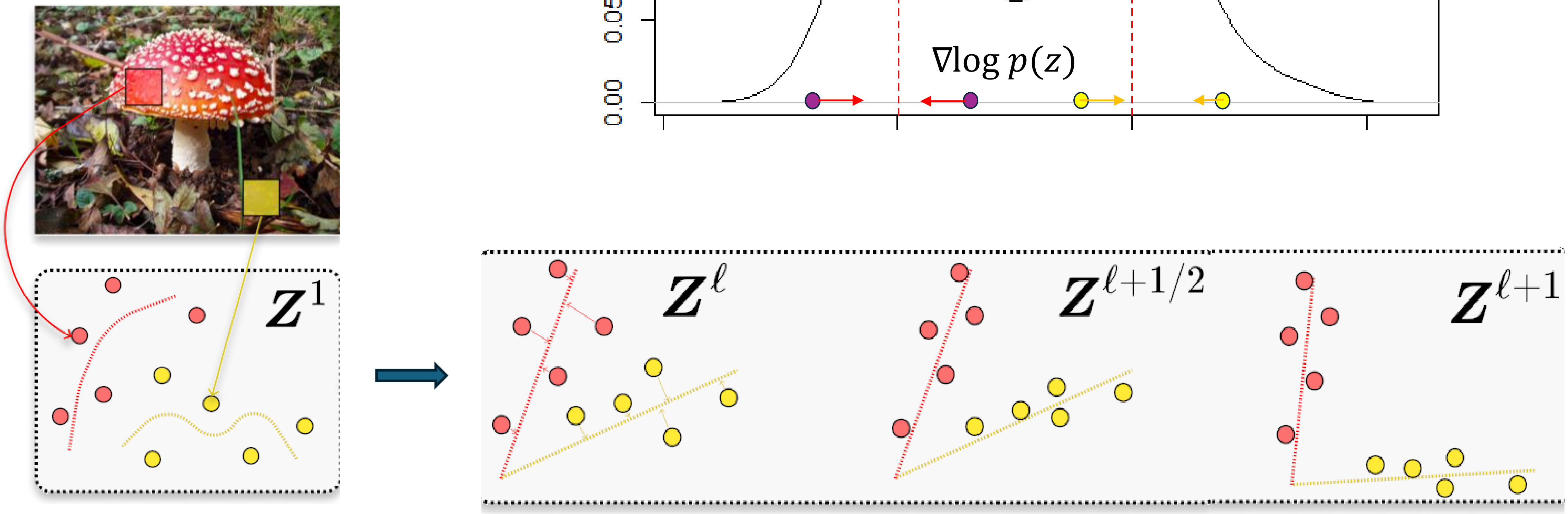




What to Learn?

low-dimensional probabilistic distributions





What to Learn?

Properties of low-dimensional structures:

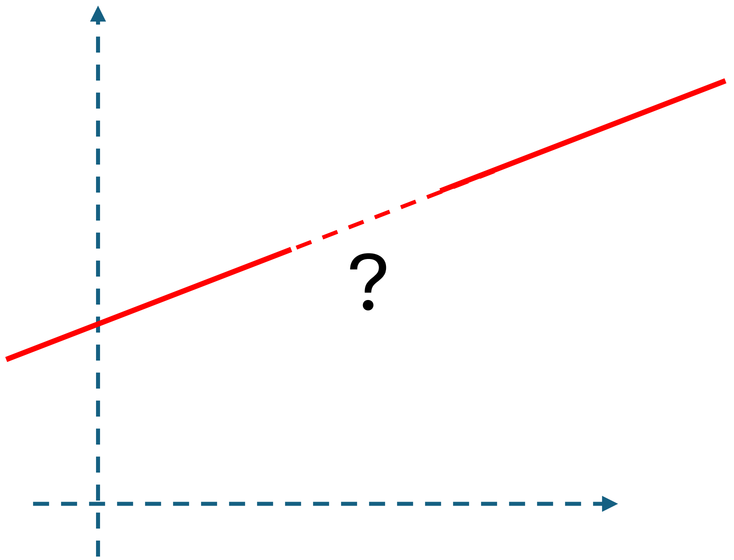
1. Completion

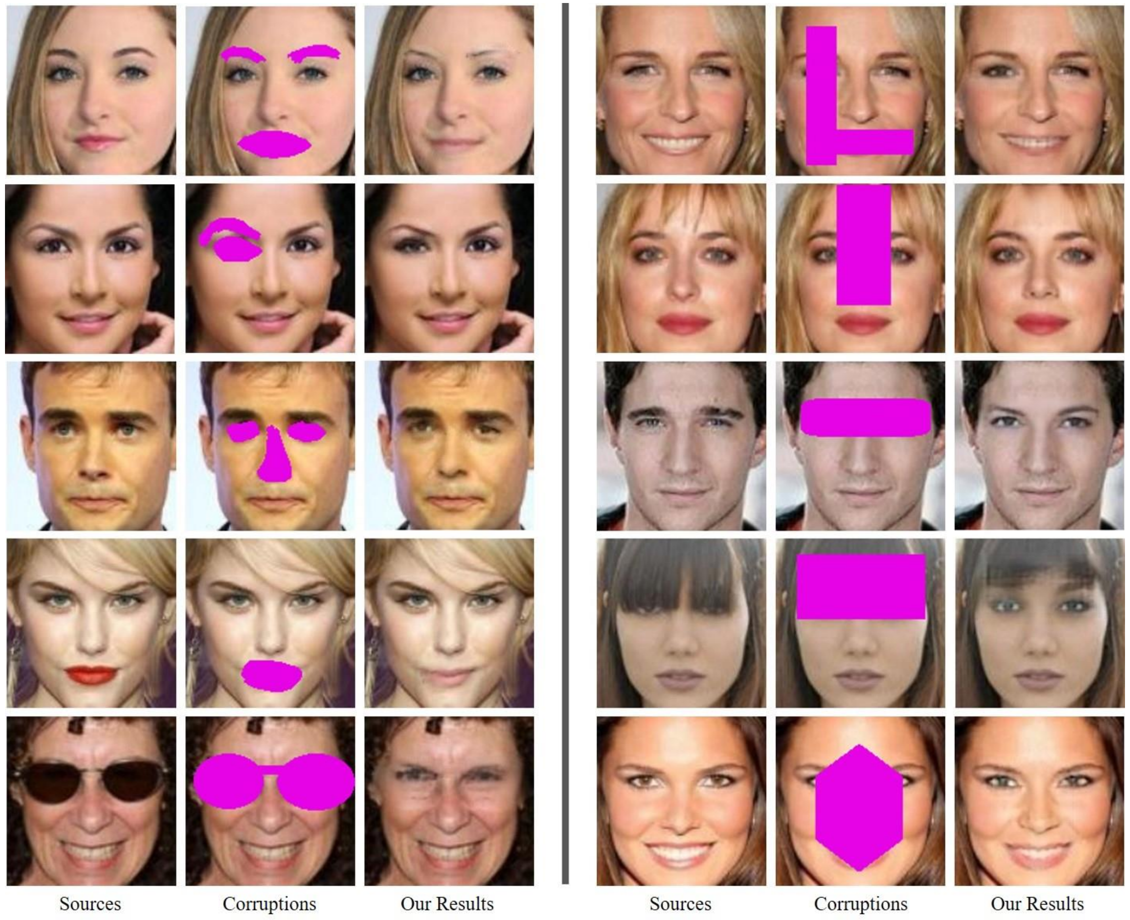
2. Denoising

3. Error Correction

What to Learn?

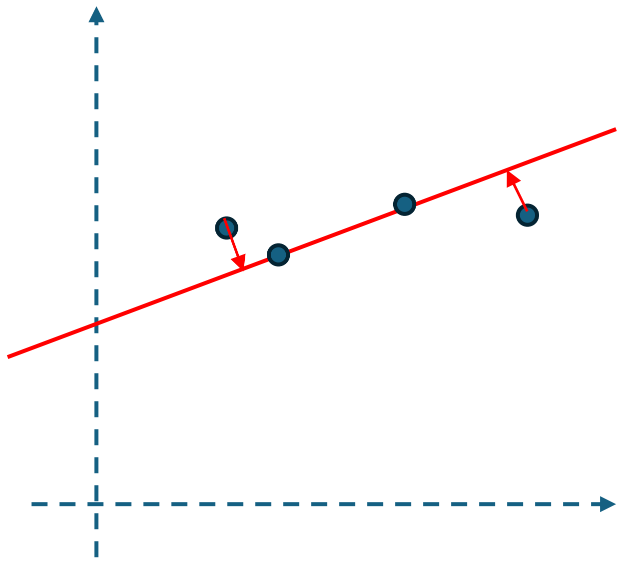
Properties of low-dimensional structures: Completion

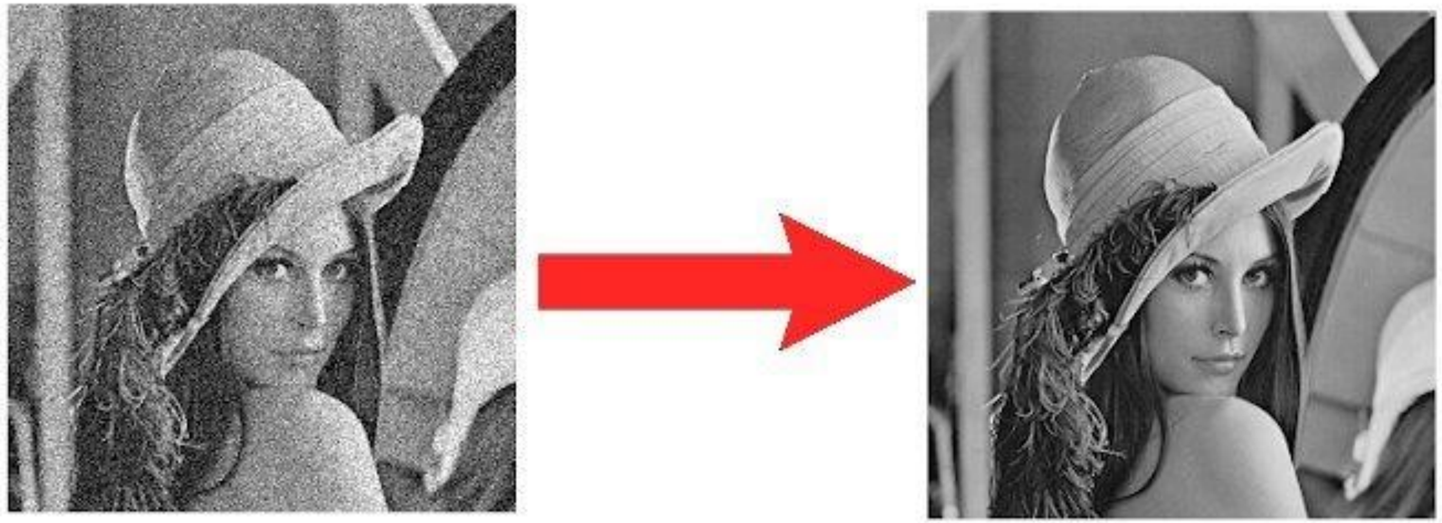




What to Learn?

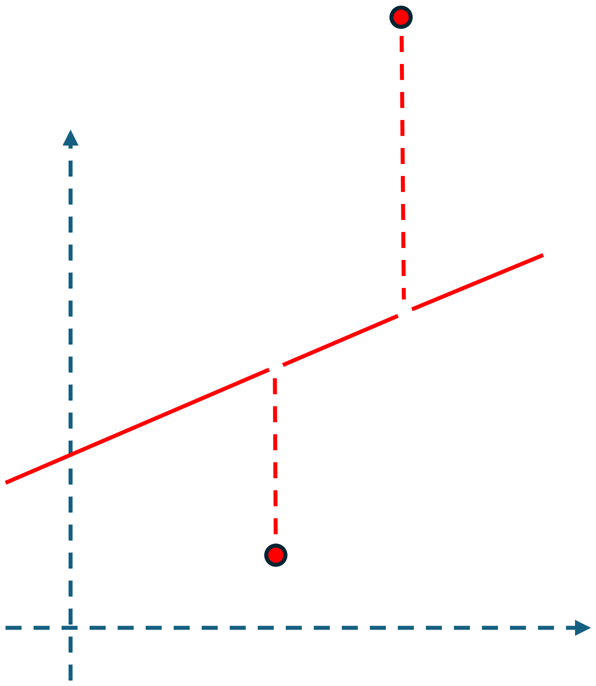
Properties of low-dimensional structures: Denoising

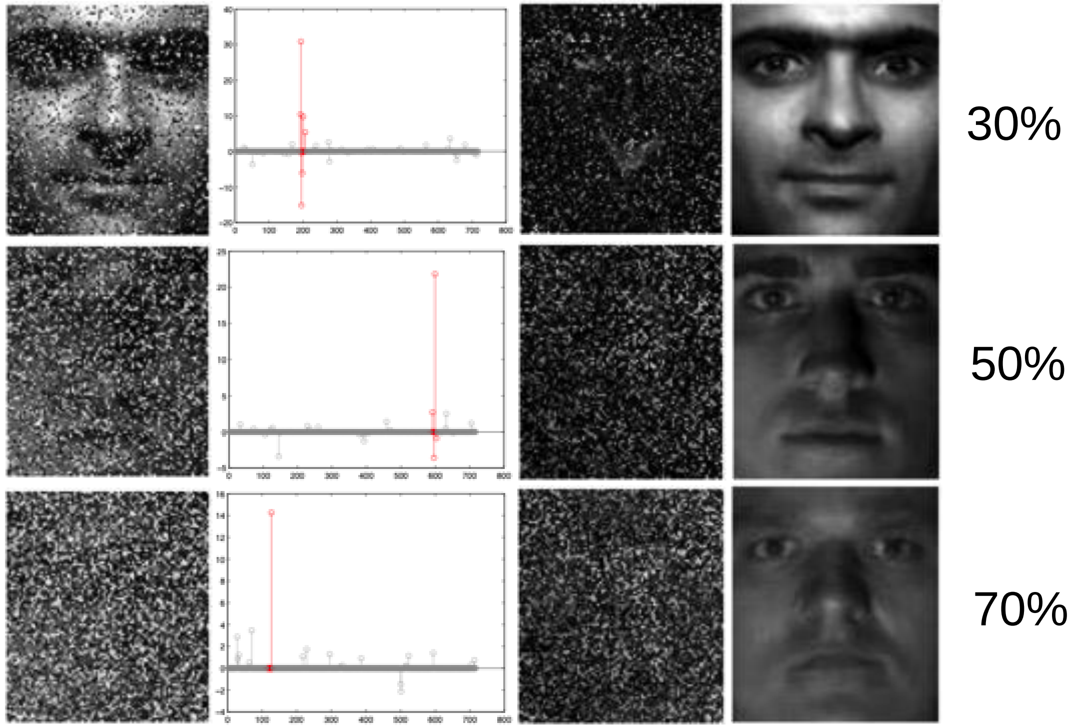


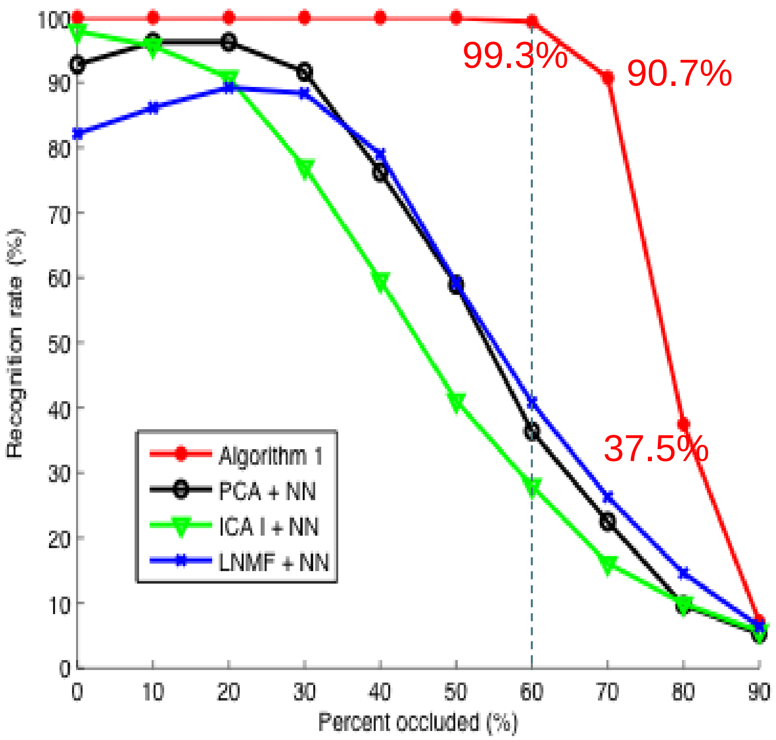


What to Learn?

Properties of low-dimensional structures: Error correction

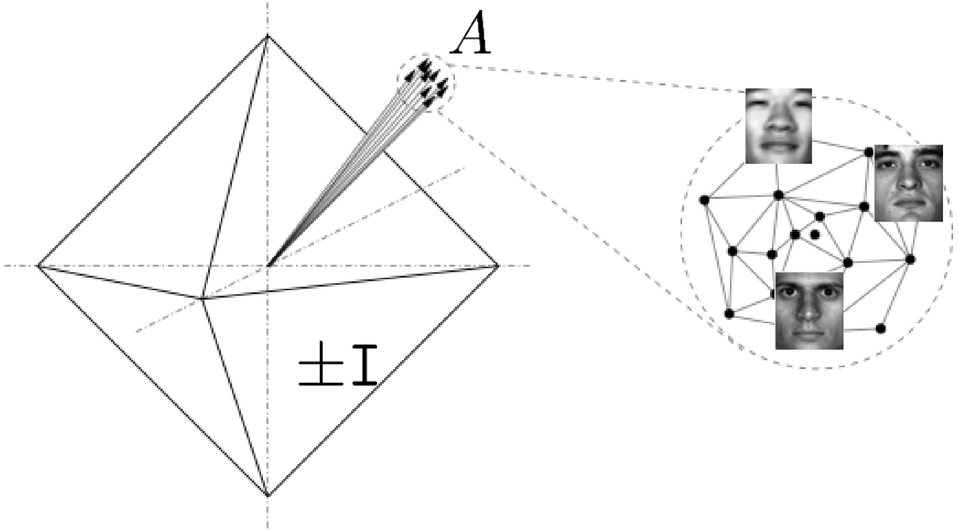






What to Learn?

Properties of low-dimensional structures: Error correction



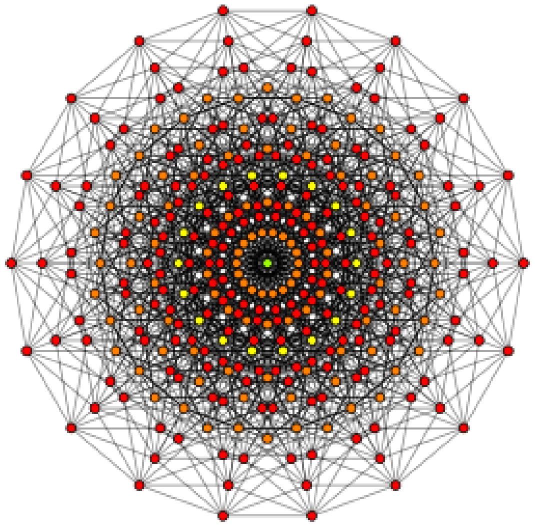
Highly coherent

(volume $\leq 1.5 \times 10^{-229}$)

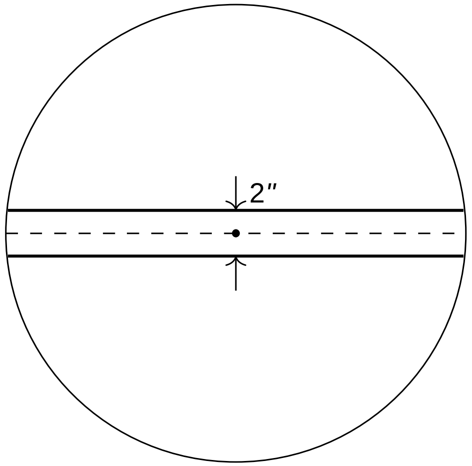
What to Learn?

Completely different geometric and statistical phenomena in high-dimensional spaces

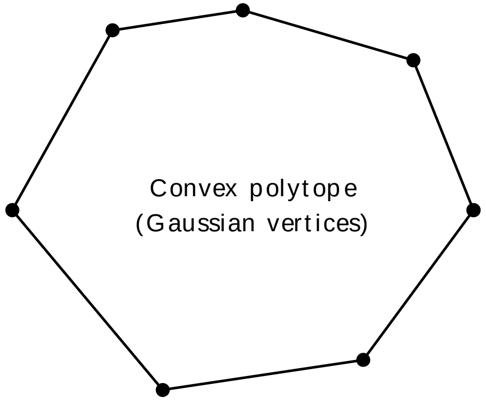
a 9-dimensional cube



a high-dim sphere



a high-dim Gaussian



How to Learn?