CS 6241: Numerical Methods for Data Science

01/22/2019

Numerical Methods:

- algo that we put on a computer to "solve a problem"
- usually involves floating point cont. math + discrete math (PageRank)

Data Science

- ??
- analysis of some information that has already been collected

This class:

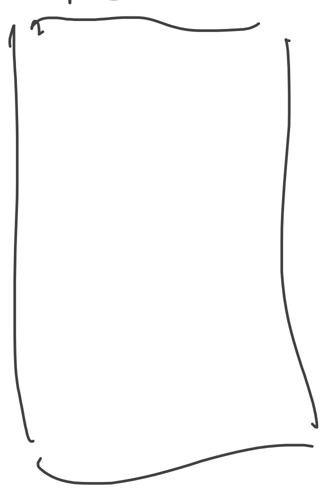
(1) Matrix/tensors

(2) Network Science

Matrices

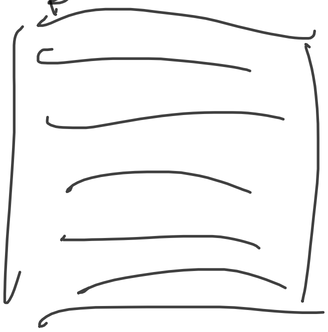
features

samples

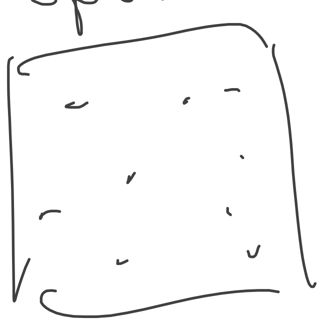


samples	feats	prob
users on FB	# friends, freq. of visits	serve ads
patients	prescriptions	corr. of prescr

Dense



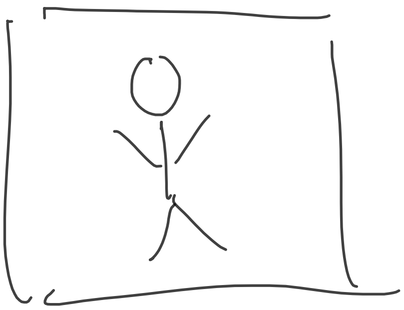
Sparse



sparse $\Longleftrightarrow$ zero pattern

helps algos

image



$X =$ 

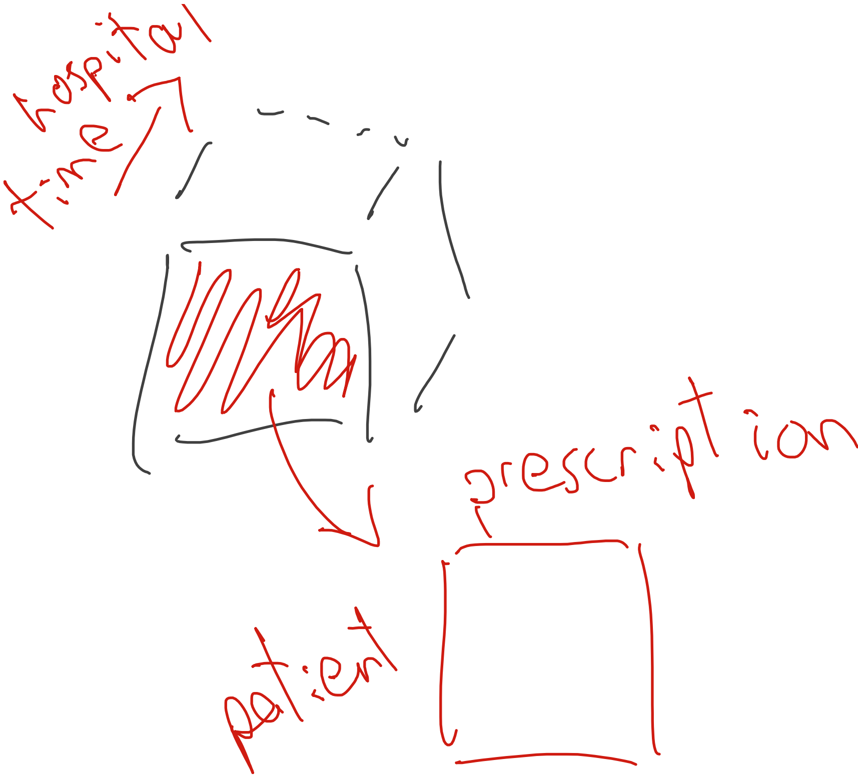
$$\frac{1}{n-1}\bar{X}^T\bar{X}$$

$$\hat{\Sigma} = \bar{X}^T = X^T\left(I - \frac{1}{n}\mathbb{1}\mathbb{1}^T\right)$$

Tensors are multi-dimensional array

(matrices are second-order tensors

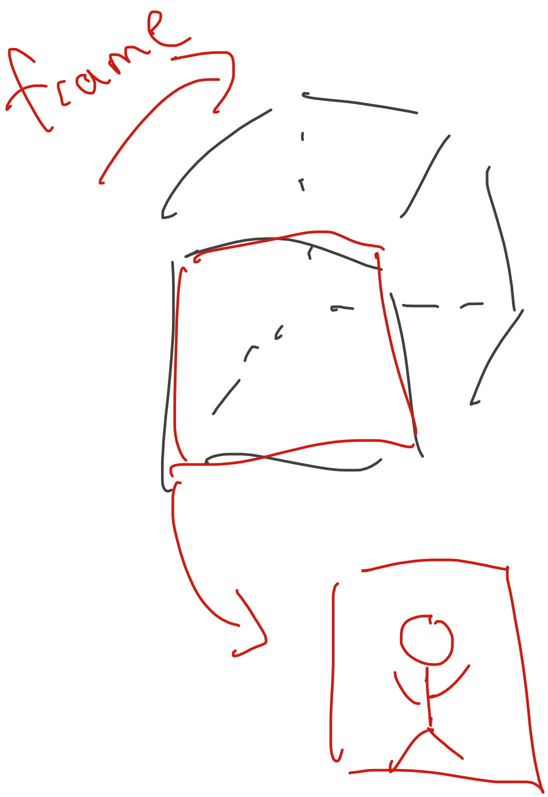
third-order tensor



sample

x feature

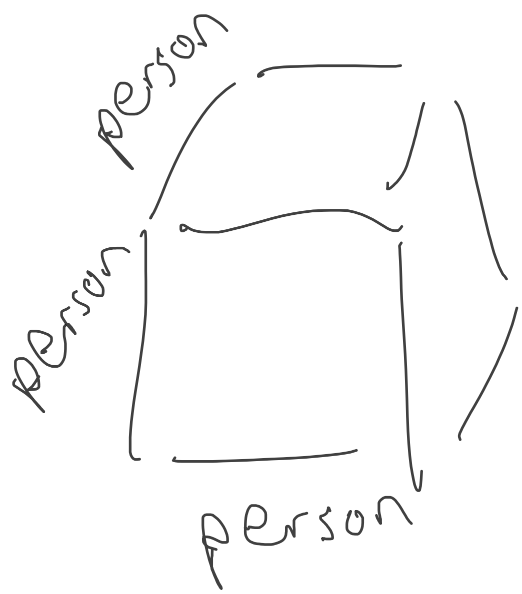
x time



movie

video

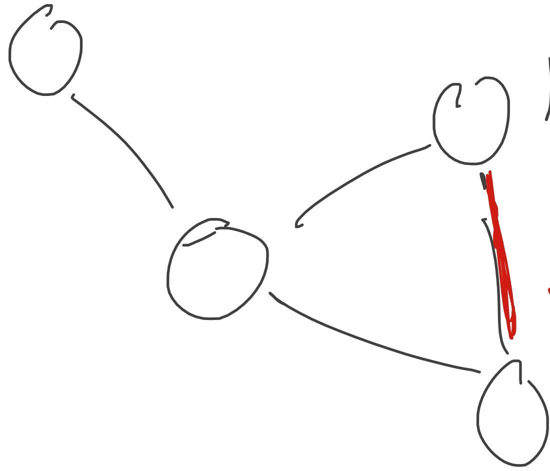
Social Network perception
(David Krackhardt)



$$\underline{A}_{ijk} = \begin{cases} 1 & \text{person } k \text{ thinks that } i \,\&\, j \text{ are friends} \\ 0 & \text{o/w} \end{cases}$$

Network analysis / science
Applied graph theory

Example: social network analysis



Node (entity)
(person)

edges (connection)
(friendship)

graphs $\Longleftrightarrow$ networks

Ecology: nodes species
edges who-eats-whom

Roads: cities
highways

Citation papers
citations

Flight

airports
direct flights

Graph $\Longleftrightarrow$ Matrix

$$A_{ij} = \begin{cases} 1 & \text{if } i \text{ connects to } j \\ 0 & \text{o/w} \end{cases}$$

$$A^{k}_{ij}$$

Networks can also be represented by Tensors

$$A_{ijk} = \begin{cases} 1 & \text{if tags } i, j, k \text{ appe} \\ & \text{in a questio} \\ 0 & \text{o/w} \end{cases}$$

Administrivia

cs.cornell.edu/courses/cs6241/2019sp

CMS, piazza

Coursework

- 2 HWs (10% each)
- Reaction paper (20%)
- Project (60%)
  - proposal (15%)
  - progress report (15%)
  - final report (30%)

teams 1, 2, 3

1. math
2. code
3. data analysis

Optimization

Nocedal + Wright

$$\min f(x)$$

subject to $x \in \Omega$

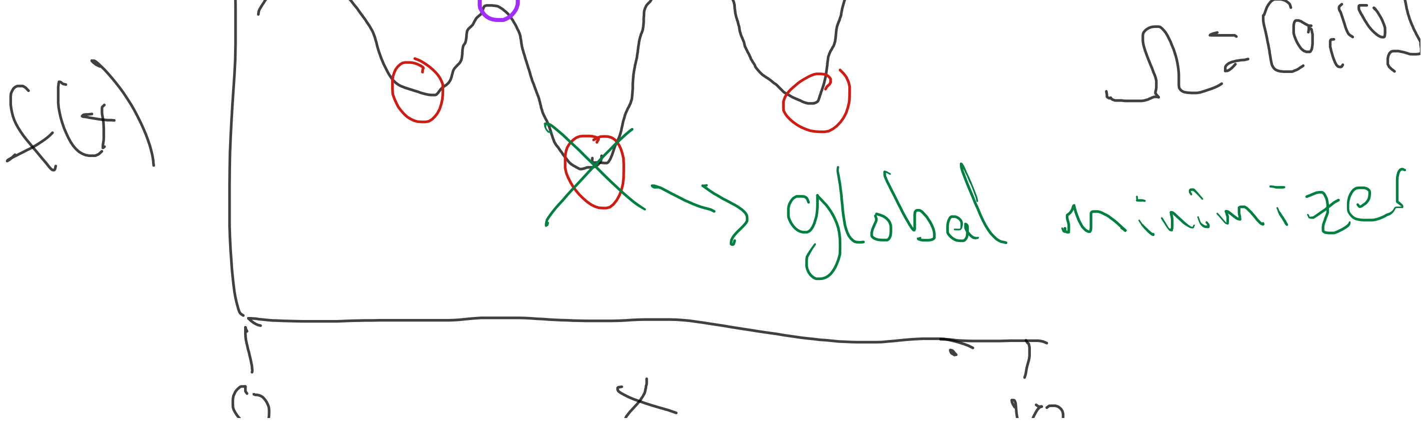
$\Omega \subseteq \mathbb{R}^n$

$f : \mathbb{R}^n \to \mathbb{R}$

f usually continuous

Ω nice in some way

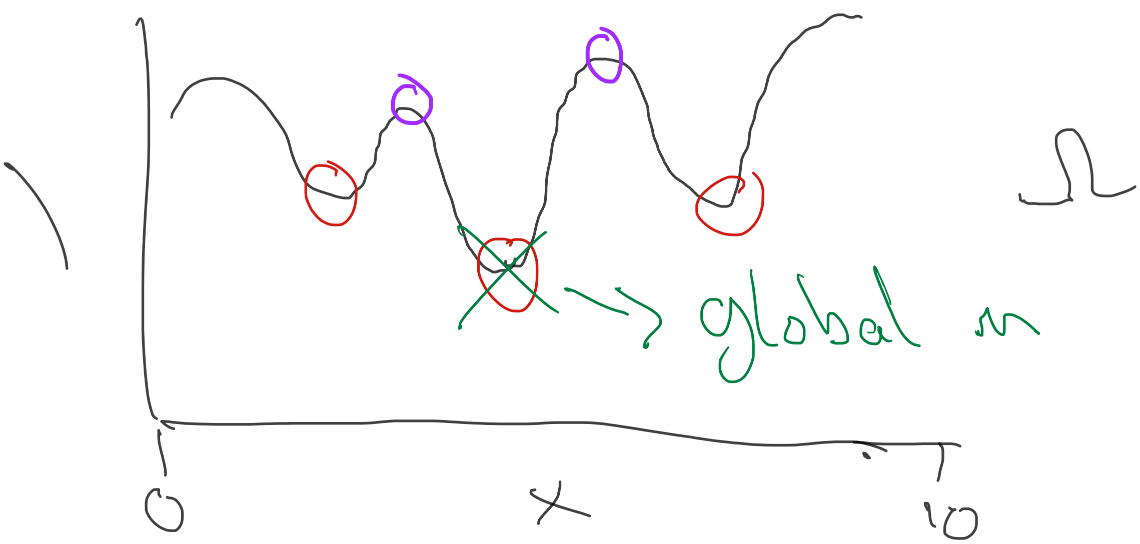
$n=1$



How do we know we are at a local minimizer x^* ?

Necessary condition:
(unconstrained)

$$\nabla f(x^*) = 0$$



$H(x)$ = Hessian at x

$$[H(x)]_{ij} = \frac{\partial^2}{\partial x_i \partial x_j} f(x)$$

Suppose H is cont. near x^*

$$\nabla f(x^*) = 0$$

$H(x^*)$ positive def.

$$\forall y \quad y^T H(x^*) y > 0$$

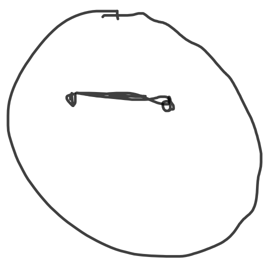
then x^* is a local minimizer

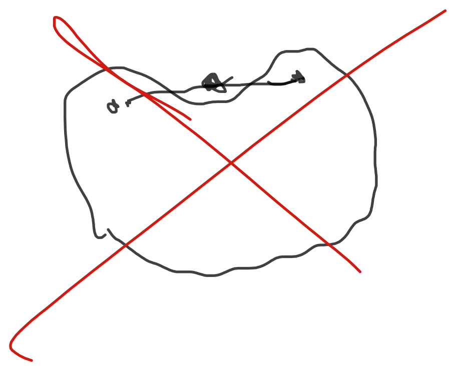
Structure: convexity

$$\min f(x)$$
$$\text{s.t. } x \in \Omega$$

convex if $\forall x, y \in \Omega$, $\alpha \in (0,1)$

$$\alpha x + (1-\alpha) y \in \Omega$$





$$f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y)$$

Example: $\min \frac{1}{2}\|x\|_2^2$

$$\text{s.t. } Ax = b$$

Convex $\Rightarrow$ every local minimum is a global minimum

Poly time global minimizers

- least squares (Thurs)
- linear programs
- SDPs

Doesn't always help! copositive matrices

$$\Omega^0 = \{ M \in \mathbb{R}^{n \times n} \mid x^T M x \geq 0, \; x \geq 0 \}$$

$$\min \; \|M\|_F^2$$
$$\text{s.t. } M \in \Omega^0$$

Numerical Linear Algebra

↳ notes on web site
Trefethen & Bau

1. Norms $\|x\|_2$
2. Eigenvalues $Ax = \lambda x$
   $\lambda \neq 0$
3. $Ax = b$

known

A sym $\Rightarrow$ $A = V \Lambda V^T$

$V^T V = I$

$\Lambda = \begin{pmatrix} \ddots & & \\ & & \\ & & \ddots \end{pmatrix}$

$AV = \Lambda V$

Any matrix A has SVD

$A = U \Sigma V^T$

$U^T U = I = V^T V$

$\Sigma = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix}$ $\quad \sigma_i \geq 0$