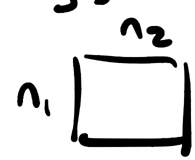
# Mini-Course on Random Tensor Models
## (and related topics)

## I. Intro and Motivation

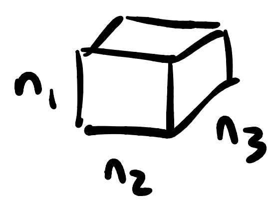
### Tensors

Order-1 tensor: vector $u \in \mathbb{R}^n$ $\quad u = (u_i)$, $\quad i \in [n] = \{1, 2, \ldots, n\}$


Order-2 tensor: matrix $M \in \mathbb{R}^{n_1 \times n_2}$ $\quad M = (M_{ij})$




Order-3 tensor: $T \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ $\quad T = (T_{ijk})$



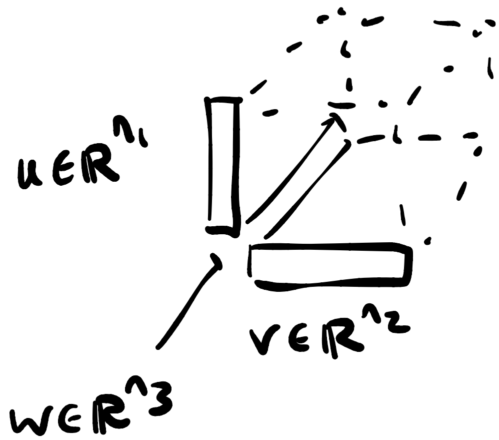

Symmetric matrix: $M_{ij} = M_{ji}$ $\quad n \times n$

Symmetric 3-tensor $T_{ijk} = T_{ikj} = T_{jki} = \cdots$

any permutation

Rank-1 matrix: $M = uv^T$ $\quad n_1 \times n_2$ $\quad M_{ij} = u_i v_j$


**Rank-1 tensor:** $T = u \otimes v \otimes w$ $\quad n_1 \times n_2 \times n_3$




$T_{ijk} = u_i v_j w_k$

__Def__ The (CP) __rank__ of a tensor $T$ is the minimum $r$ such that $T$ can be expressed as the sum of $r$ rank-1 tensors.

Order 3: $$T = \sum_{i=1}^{r} u^{(i)} \otimes v^{(i)} \otimes w^{(i)}$$

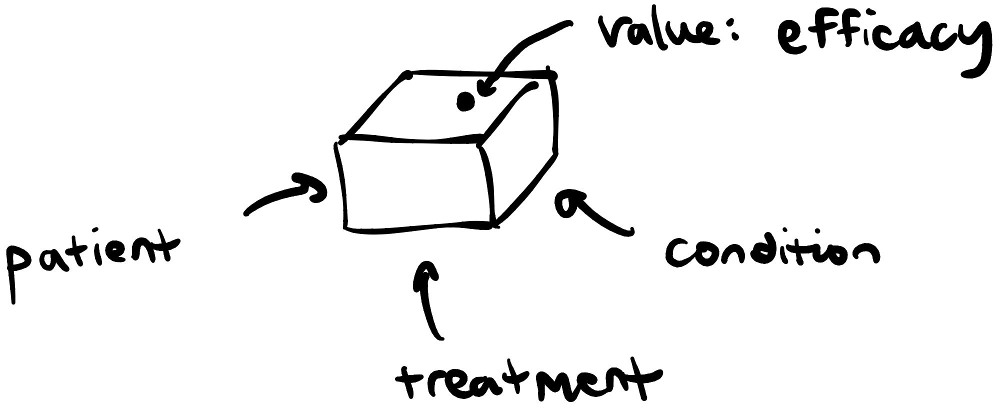
$u^{(i)} \in \mathbb{R}^{n_1}$

Order 2: coincides with matrix rank

Why tensors? in statistics, data science, ML, ...

Two reasons:

1) Multi-way data




* Might expect/hope data is (approx) low rank

* Want to learn underlying structure (like PCA)
  ↳ low-rank approx

<u>Model 1</u> (Tensor PCA / Spiked Tensor)

Observe $T = \lambda v^{\otimes 3} + Z$ $\quad n \times n \times n$

$Z$: iid $N(0,1)$ tensor

$\lambda \geq 0$
SNR

$v^{\otimes 3} = v \otimes v \otimes v$

$v \in \mathbb{R}^n$ – planted signal, unknown

$\|v\| = \sqrt{n}$

Goal: recover/estimate $v$

Note: order-2 case is spiked Wigner

Objective: want algorithm that is

(i) statistically efficient (small $\lambda$)

(ii) computationally efficient (fast runtime)

polynomial time

Spoiler: can't have both!

Focus: computational threshold

2) Method of moments

Goal: learn a distribution $D$ over $\mathbb{R}^n$ from $y_1, \ldots, y_N \overset{iid}{\sim} D$

$y_i \in \mathbb{R}^n$

From samples, can estimate

1st moment: $\mathbb{E}_{y \sim D}[y]$

$2^{nd}$ moment: $\mathbb{E}[yy^T]$ for $\mathbb{E}[y]=0$, covariance matrix

$3^{rd}$ moment: $\mathbb{E}[y^{\otimes 3}]$ $n \times n \times n$

$\vdots$

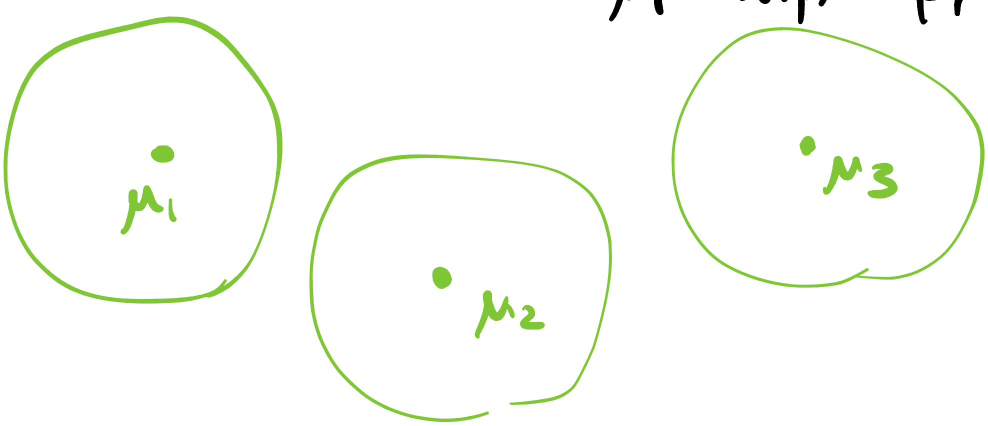
<u>Ex</u> (Gaussian mixture)

Unknown centers $\mu_1, \mu_2, \dots, \mu_r \in \mathbb{R}^n$

$$D: \quad y = x + z$$

$z \sim N(0, I_n)$

$$x = \begin{cases} \mu_1 & \text{w.p. } p_1 \\ \mu_2 & \text{w.p. } p_2 \\ & \vdots \\ \mu_r & \text{w.p. } p_r \end{cases}$$




Goal: find $\mu_1, \dots, \mu_r$

$$\mathbb{E}[y] = \mathbb{E}[x] + \underbrace{\mathbb{E}[z]}_{0} = \sum_{i=1}^{r} p_i \mu_i$$

$$\mathbb{E}[yy^T] = \mathbb{E}\,(x+z)(x+z)^T = \mathbb{E}[xx^T] + \underbrace{\mathbb{E}[zz^T]}_{I}$$

$\mathbb{E}[xz^T] = 0$

$\mathbb{E}[zx^T] = 0$

$$= I + \sum_{i=1}^{r} p_i \mu_i \mu_i^T$$

$$\overset{\text{learn}}{\rightsquigarrow} \sum_{i=1}^{r} p_i \mu_i \mu_i^T$$

like eigendecomp but $\mu_i$ not orthog.

* First 2 moments are not enough

$$\mathbb{E}[y^{\otimes 3}] \overset{\text{learn}}{\rightsquigarrow} \sum_{i=1}^{r} p_i \mu_i^{\otimes 3}$$

* 3rd moment is (sometimes) enough info

↳ $r$ small enough

Model 2 (Tensor Decomposition)

Given $T = \sum_{i=1}^{r} (u^{(i)})^{\otimes 3}$ $\quad u^{(i)} \in \mathbb{R}^n$

Goal: recover $\{u^{(1)}, \ldots, u^{(r)}\}$

un-ordered

## Roadmap

- Solution(s) to tensor PCA & tensor decomp

* Challenge: most tensor problems are NP-hard

worst-case

- compute rank [Håstad '90]
- best low-rank approx [Hillar, Lim '09]

⋮

whereas matrices have SVD, eigendecomp, ...