Aside: Cumulants

Let $m_1 = \mathbb{E}_D[X]$, $m_2 = \mathbb{E}_D[X^2]$, etc be moments

There's another basis that's sometimes more convenient

cumulants
$$\begin{cases} \kappa_1 = m_1 \\ \kappa_2 = m_2 - m_1^2 \\ \kappa_3 = 2m_1^3 - 3m_1 m_2 + m_3 \\ \vdots \end{cases}$$

Fact: If $X \perp\!\!\!\perp Y$ then $\kappa_i(X+Y) = \kappa_i(X) + \kappa_i(Y)$

Not true for moments!

Easier way to think about cumulants

$$\kappa(t) = \log \mathbb{E}[e^{tX}]$$
$$= \sum_{i=1}^{\infty} \kappa_i \frac{t^i}{i!}$$

Now the fact is easy to prove

$$\log \mathbb{E}[e^{t(X+Y)}] = \log\left(\mathbb{E}[e^{tX}]\,\mathbb{E}[e^{tY}]\right)$$

$$\Rightarrow \sum_{i=1}^{\infty} \kappa_i(X+Y)\frac{t^i}{i!} = \sum_{i=1}^{\infty} \kappa_i(X)\frac{t^i}{i!} + \sum_{i=1}^{\infty} \kappa_i(Y)\frac{t^i}{i!}$$

Many strange identities that come up when we use tensor decompositions can be understood thru power series

Independent Component Analysis

Given samples $y = Ax + b$ ← Gaussian noise

full rank independent coordinates

want to learn A

Claim: If each $x_i \sim N(0,1)$, it's impossible

This is because both

orthogonal

$$y = Ax + b \quad \text{and} \quad y' = AUx + b$$

generate the same distribution (i.e. check $A = I$)

Meta Theorem: If the x_i's are nongaussian, there are efficient algorithms to learn A up to permutation and scaling of columns

Can we use tensor decompositions?

For now, let's assume $b = 0$

Then we can try $\mathbb{E}[y^{\otimes 4}]$

$$= \mathbb{E}\left[(x_1 A_1 + \cdots x_n A_n)^{\otimes 4}\right]$$

$$= \sum_{ijkl} \underbrace{\mathbb{E}[x_i x_j x_k x_l]}_{D} A_i \otimes A_j \otimes A_k \otimes A_l$$

Is D diagonal? If it was, could use Jennrich

Unfortunately not, consider

$$D_{iijj} = \mathbb{E}[x_i^2 x_j^2] > 0$$

Can we correct D to make it diagonal, by subtracting off expressions involving lower order moments?

Digression: Multivariate cumulants

$$K(\vec{t}) = \log \mathbb{E}\left[e^{\vec{t}^T \vec{x}}\right]$$

$$= \sum_{i=1}^{\infty} \frac{\langle K_i, \vec{t}^{\otimes i} \rangle}{i!}$$

K_1 is a vector, K_2 is a matrix, ... K_i is an i^{th} order tensor

Claim: If x_i's are $\perp\!\!\!\perp$ then K_j's are diagonal

Proof: Consider the case $n=2$ and by $\perp\!\!\!\perp$

$$\underbrace{\kappa_2(X_1+X_2)}_{\kappa_2(\vec{t})\big|_{\vec{t}=e_1+e_2}} = \underbrace{\kappa_2(X_1)}_{\kappa_2(\vec{t})\big|_{\vec{t}=e_1}} + \kappa_2(X_2)$$

etc

$$\Rightarrow (e_1+e_2)^T \kappa_2 (e_1+e_2) = e_1^T \kappa_2 e_1 + e_2^T \kappa_2 e_2$$

But this can only happen if κ_2 is diagonal. $\boxtimes$

So the cumulants, whatever they are, give us a way to fix the moments

Note: Many works use cumulants w/o calling them that, can make identities look mysterious

Fact: The third and higher cumulants of a Gaussian are zero

So if X_i's have nonzero cumulants:

cumulants + Jennrich = ICA