Independence ~~& Correlations~~

Jonathan Pillow

Mathematical Tools for Neuroscience (NEU 314)
Fall, 2021

lecture 15

PCA warmup problem

the data

$$X = \overbrace{\begin{bmatrix} - & \vec{x}_1 & - \\ - & \vec{x}_2 & - \\ & \vdots & \\ - & \vec{x}_N & - \end{bmatrix}}^{d} \Bigg\} N$$

collection of N data vectors

As a reminder from last week's quiz, PCA involves the following two steps:

1. Compute the ("covariance" or "2nd moment") matrix: $C = X^\top X$
2. Perform SVD: $C = USU^\top$

- the top k principal components (PCs) are the first k columns of U!

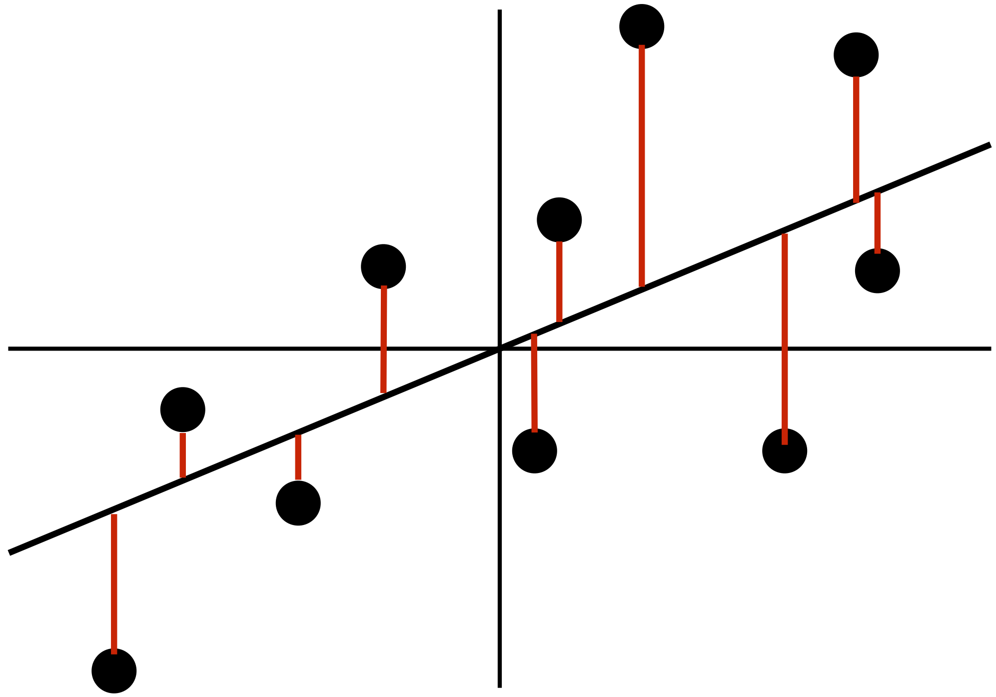
Question: let us denote the SVD of X by: $X = U_x S_x V_x^T$

What is the relationship between the SVD of C and the SVD of X?
(That is, what is the relationship between U, S and U_x, S_x, V_x, if any?)

- put another way, is there a way to get the PCs & their fraction-of-variance explained without computing C?

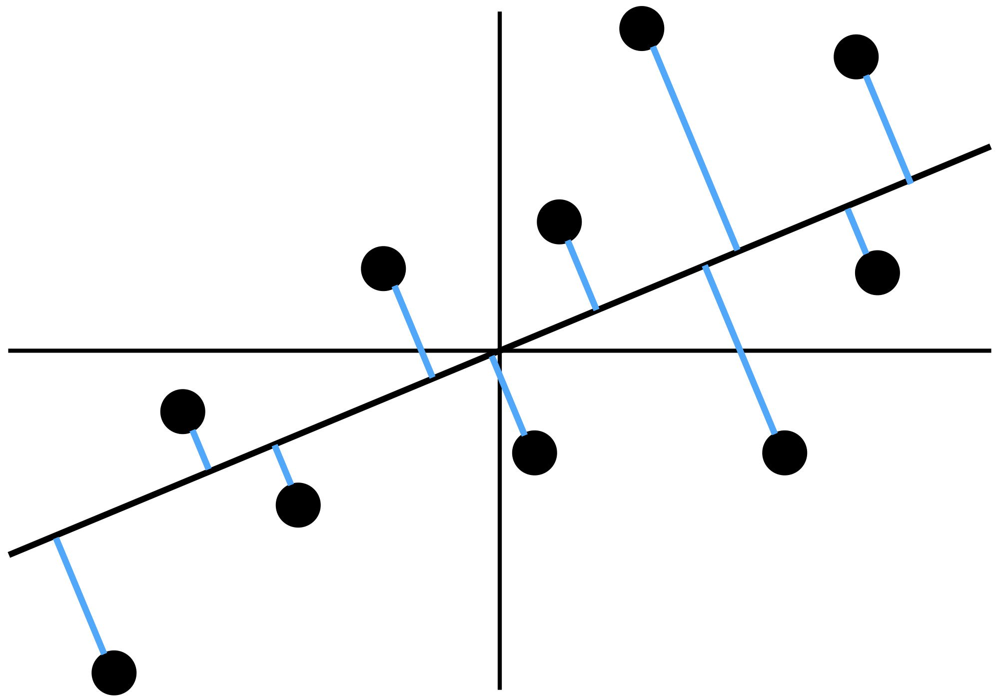
warmup problem: PCA & regression

Consider the following two kinds of distances from datapoints to a line:



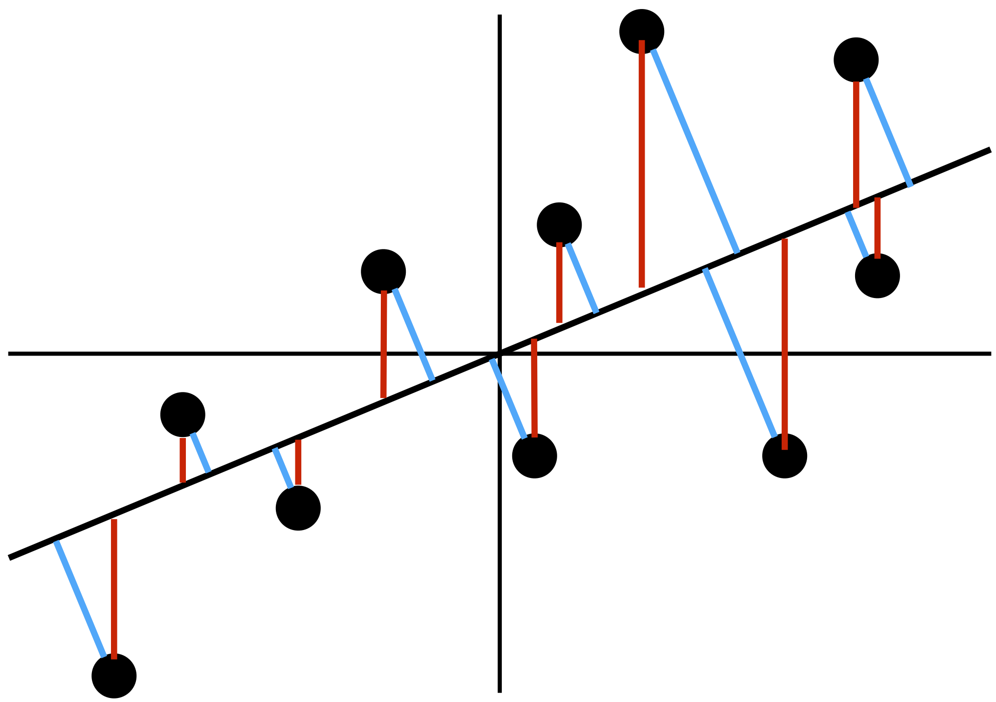
warmup problem: PCA & regression

Consider the following two kinds of distances from datapoints to a line:



warmup problem: PCA & regression

Consider the following two kinds of distances from datapoints to a line:



1. Which is correct:

a) PCA minimizes sum of red^2 and Least Squares minimizes sum of blue^2
b) PCA minimizes sum of blue^2 and Least Squares minimizes red^2
c) PCA and Least Squares both minimize sum of blue^2
d) PCA and Least Squares both minimize sum of red^2

2. There is a special name for the red lines, what is it?

practice problems: probability

P(x,y)

y \ x	1	2	3
3	0.1	0	0.1
2	0	0.3	0
1	0.2	0.1	0.2

1. Compute the marginal P(x)
2. Compute the marginal P(y)
3. Compute the conditional P(y | x = 2)
4. Compute the conditional P(x | y = 1)
5. What is the most probable value for y?

practice problems: probability

P(x,y)

y \ x	1	2	3
3	0.1	0	0.1
2	0	0.3	0
1	0.2	0.1	0.2

6. What is the conditional P(x | y > 1)?
7. What is the conditional P(x | x = y)?
8. What is the conditional expectation $\mathbb{E}[x \,|\, y = 3]$

<u>bonus</u>: compute the conditional variance $\mathrm{var}(x \,|\, y = 3)$

practice problems: Bayes rule

Consider the following model describing how a single neuron responds to houses and faces, which is given by a pair of conditional distributions:

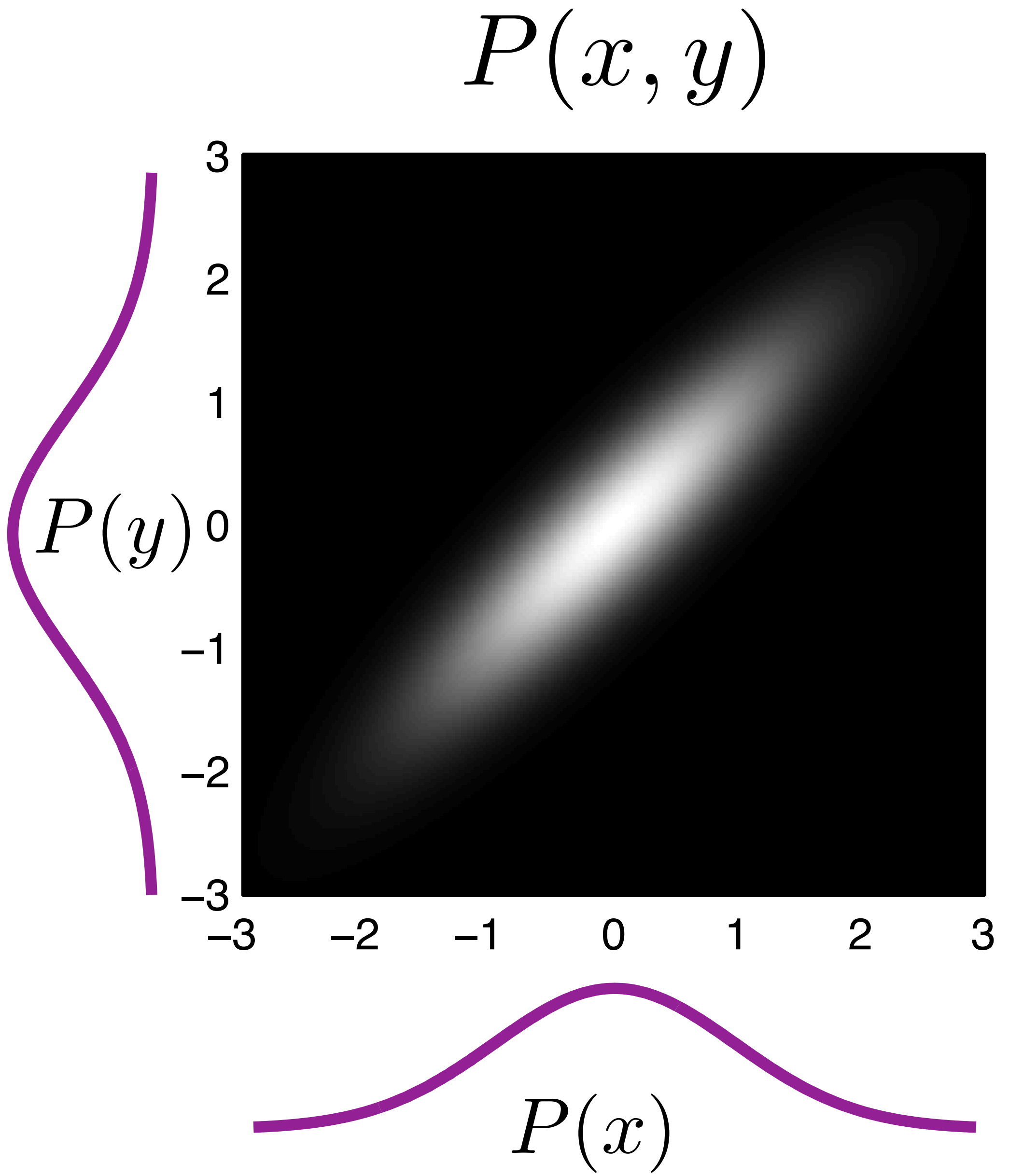
# spikes	0	1	2	3	4
P(# spikes I "house")	0.1	0.1	0.3	0.4	0.1
P(# spikes I "face")	0.1	0.4	0.3	0.2	0

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)}$$

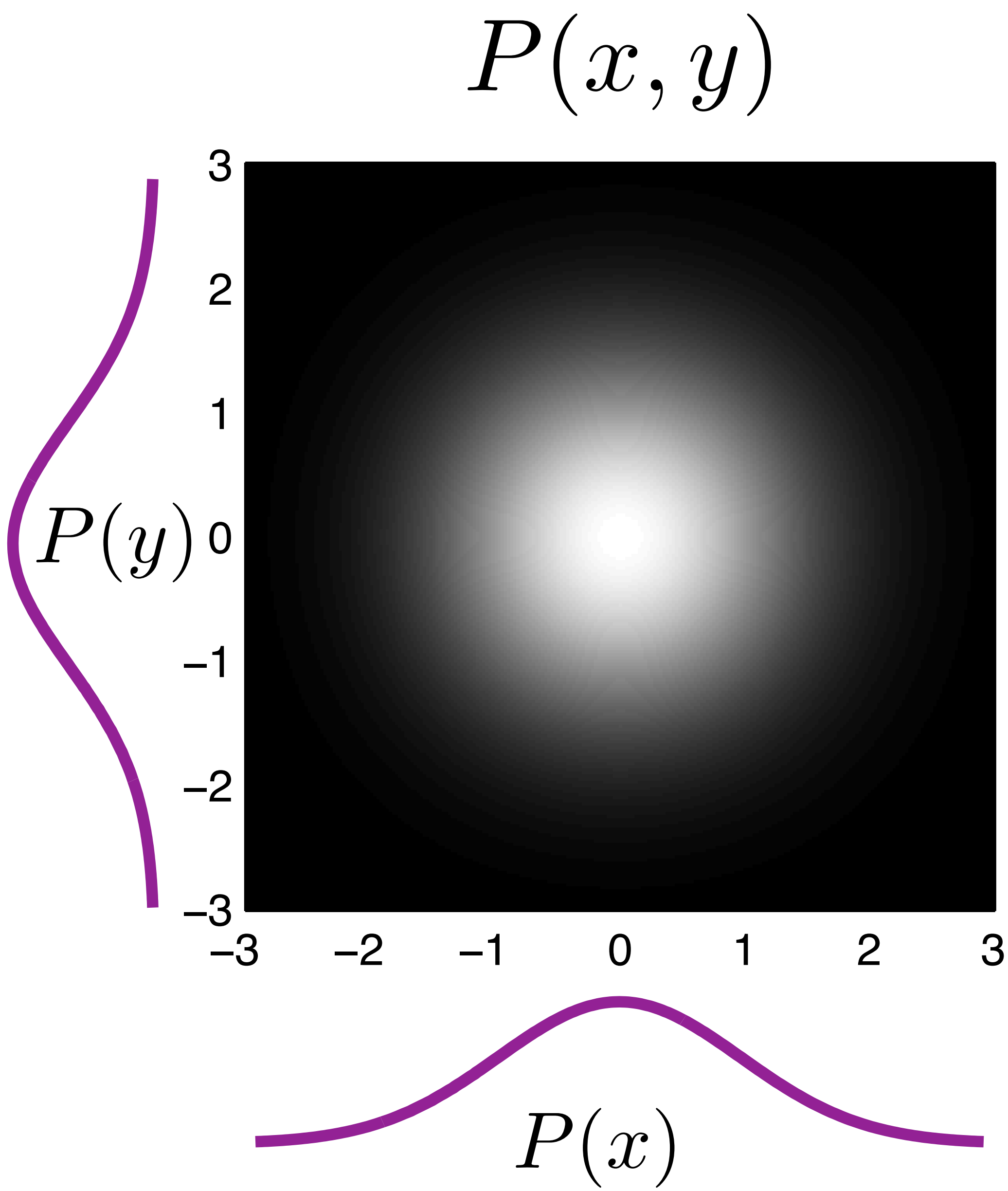
Furthermore, suppose P(house) = P(face) = 0.5

1) Does this neuron respond more to houses or faces?
2) What is the most probable stimulus if you observe 2 spikes?
   • Compute P(face I 2 spikes) and P(house I 2 spikes)
3) What is the most probable stimulus if you observe 3 spikes?
   • Compute P(face I 3 spikes) and P(house I 3 spikes)
4) Is there any response for which you can be certain of what the stimulus was?
5) Re-answer #3 under the prior that P(house) = 0.2 , P(face) = 0.8

Joint Distribution



independence



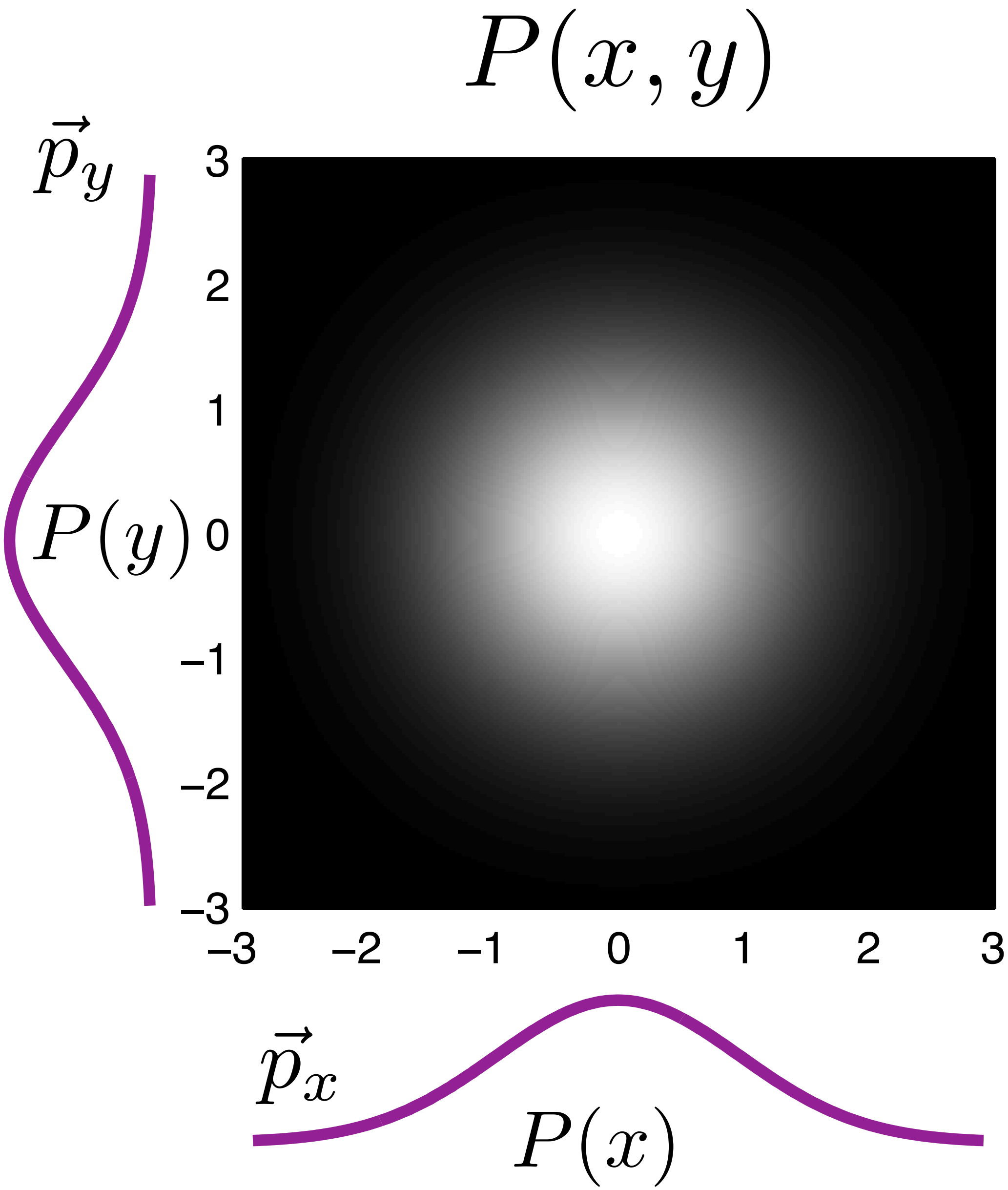
Definition: *x*, y are *independent* iff ("if and only if")

$$P(x, y) = P(x)P(y)$$

independence

Definition: *x*, y are *independent* iff
("if and only if")

$$P(x,y) = P(x)P(y)$$



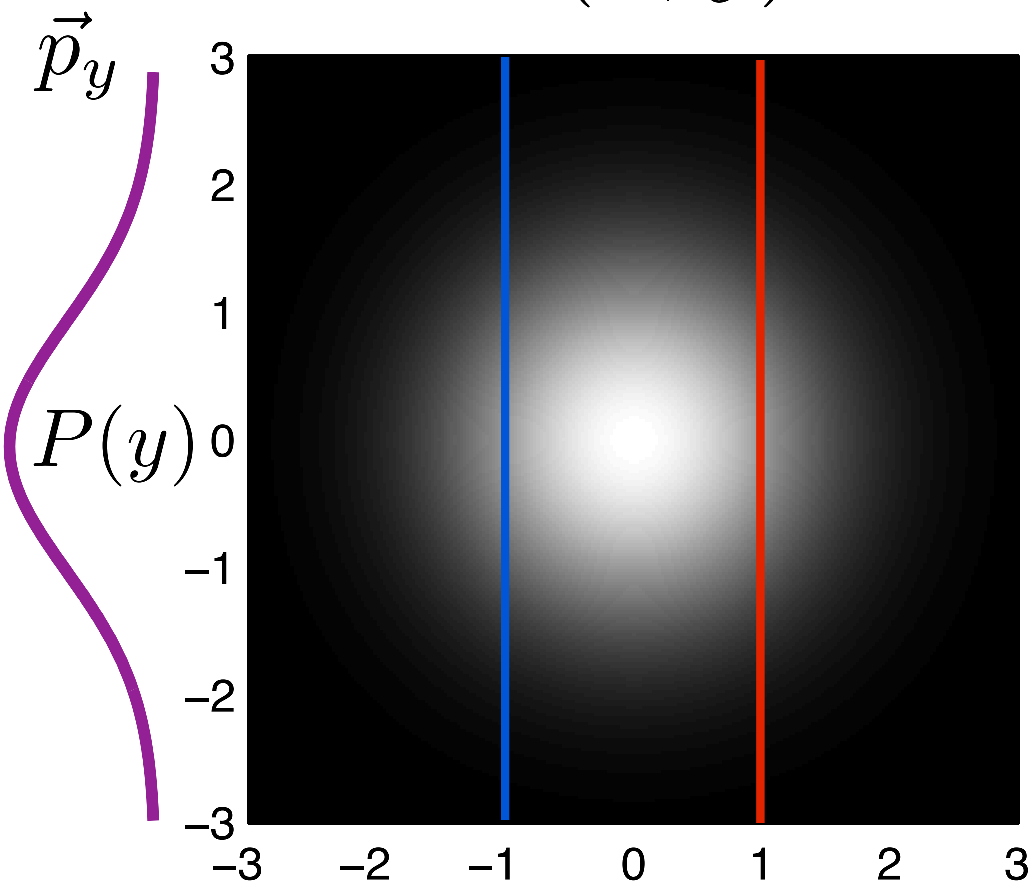
In linear algebra terms:

$$P(x,y) = \vec{p}_y \, \vec{p}_x^{\,T}$$

(outer product)

independence

$$P(x, y)$$



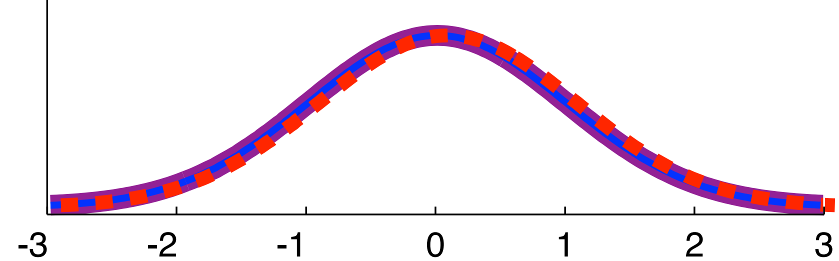
Original definition:

$$P(x, y) = P(x)P(y)$$

Equivalent definition:

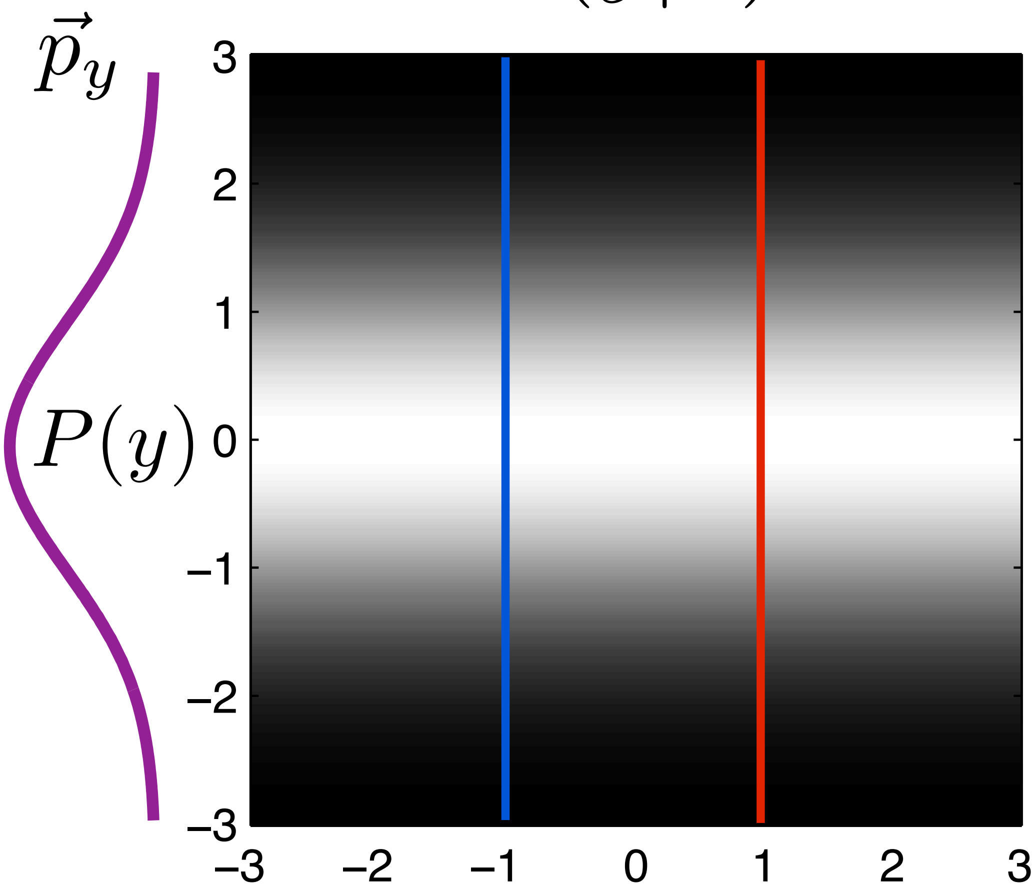
$$P(y|x) = P(y)$$ for all x

All conditionals are the same!



independence

$$P(y|x)$$



Original definition:

$$P(x, y) = P(x)P(y)$$

Equivalent definition:

$$P(y|x) = P(y)$$ for all x

All conditionals are the same!