Lisa Yan
CS109

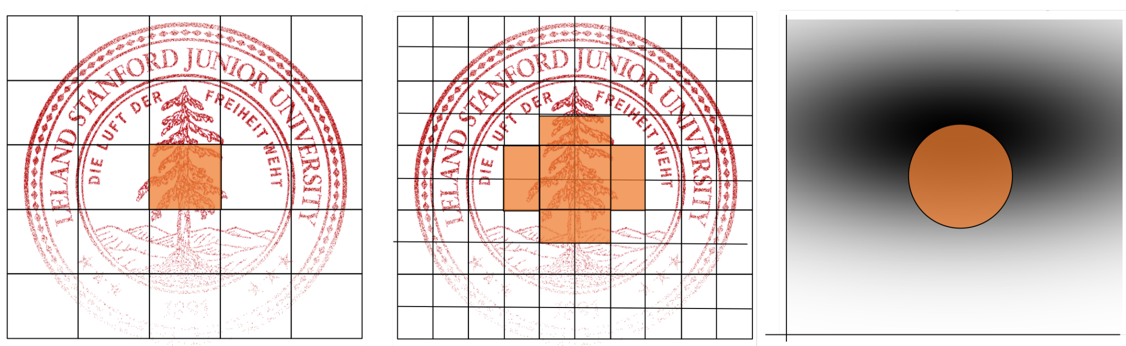
Lecture Notes #12
May 11, 2020

Continuous Joint Distributions

Based on a chapter by Chris Piech and Lisa Yan

Continuous Joint Distributions

Of course joint variables don't have to be discrete only, they can also be continuous. As an example: consider throwing darts at a dart board. Because a dart board is two dimensional, it is natural to think about the X location of the dart and the Y location of the dart as two random variables that are varying together (aka they are joint). However since x and y positions are continuous we are going to need new language to think about the likelihood of different places a dart could land. Just like in the non-joint case continuous is a little tricky because it isn't easy to think about the probability that a dart lands at a location defined to infinite precision. What is the probability that a dart lands at exactly (X=456.234231234122355, Y = 532.12344123456)?:



Lets build some intuition by first starting with discretized grids. On the left of the image above you could imagine where your dart lands is one of 25 different cells in a grid. We could reason about the probabilities now! But we have lost all nuance about how likelihood is changing within a given cell. If we make our cells smaller and smaller we eventually will get a second derivative of probability: once again a probability density function. If we integrate under this joint-density function in both the x and y dimension we will get the probability that x takes on the values in the integrated range and y takes on the values in the integrated range!

Random variables X and Y are Jointly Continuous if there exists a Probability Density Function (PDF) $f_{X,Y}$ such that:

$$P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = \int_{a_1}^{a_2} \int_{b_1}^{b_2} f_{X,Y}(x, y) dy\ dx$$

Using the PDF we can compute marginal probability densities:

$$f_X(a) = \int_{-\infty}^{\infty} f_{X,Y}(a, y) dy$$
$$f_Y(b) = \int_{-\infty}^{\infty} f_{X,Y}(x, b) dx$$

Independence with Multiple RVs (Continuous Case)

Two continuous random variables X and Y are called **independent** if:

$$P(X \leq a, Y \leq b) = P(X \leq a)P(Y \leq b) \text{ for all } a, b$$

This can be stated equivalently as:

$$F_{X,Y}(a,b) = F_X(a)F_Y(b) \text{ for all } a, b$$
$$f_{X,Y}(a,b) = f_X(a)f_Y(b) \text{ for all } a, b$$

More generally, if you can factor the joint density function, then your continuous random variables are independent:

$$f_{X,Y}(x,y) = h(x)g(y) \text{ where } -\infty < x, y < \infty$$

Bivariate Normal Distribution

Many times, we talk about multiple Normal (Gaussian) random variables, otherwise known as Multivariate Normal (Gaussian) distributions. Here, we talk about the two-dimensional case, called a Bivariate Normal Distribution. X_1 and X_2 follow a bivariate normal distribution if their joint PDF is

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)}.$$

We often write the distribution of the *vector* $\mathbf{X} = (X_1, X_2)$ as $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} = (\mu_1, \mu_2)$ is a *mean vector* and $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$ is a *covariance matrix*.

Note that ρ is the correlation between X_1 and X_2, and $\sigma_1, \sigma_2 > 0$. We defer to Ross Chapter 6, Example 5d, for the full proof, but it can be shown that the marginal distributions of X_1 and X_2 are $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$, respectively.

Example 1

Let $\mathbf{X} = (X_1, X_2) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} = (\mu_1, \mu_2)$ and $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$, a diagonal covariance matrix.

Noting that the correlation between X_1 and X_2 is $\rho = 0$:

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)} = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-(x_1-\mu_1)^2/(2\sigma_1^2)} \frac{1}{\sigma_2\sqrt{2\pi}} e^{-(x_2-\mu_2)^2/(2\sigma_2^2)}$$

In other words, for Bivariate Normal RVs, if $\mathrm{Cov}(X_1, X_2) = 0$, then X_1 and X_2 are independent. Wild!

Joint CDFs

For two random variables X and Y that are jointly distributed, the joint cumulative distribution function $F_{X,Y}$ can be defined as

$$F_{X,Y}(a, b) = P(X \leq a, Y \leq b)$$

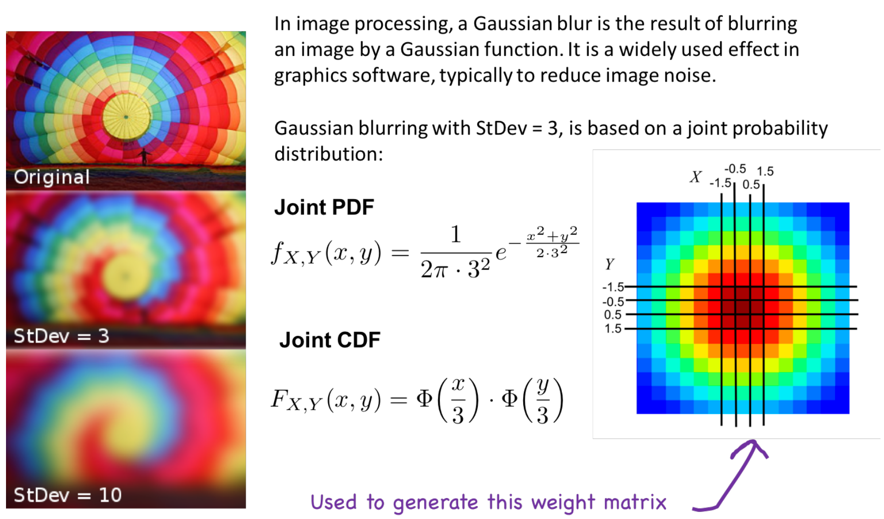
$$\begin{aligned}
F_{X,Y}(a, b) &= \sum_{x \leq a} \sum_{y \leq b} p_{X,Y}(x, y) && X, Y \text{ discrete} \\
F_{X,Y}(a, b) &= \int_{-\infty}^{a} \int_{-\infty}^{b} f_{X,Y}(x, y)\,dy\,dx && X, Y \text{ continuous} \\
f_{X,Y}(a, b) &= \frac{\partial^2}{\partial a \partial b} F_{X,Y}(a, b) && X, Y \text{ continuous}
\end{aligned}$$

It can be shown via geometry that to calculate probabilities of joint distributions, we can use the CDF as follows, for both jointly discrete and jointly continuous RVs:

$$P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = F_{X,Y}(a_2, b_2) - F_{X,Y}(a_1, b_2) - F_{X,Y}(a_2, b_1) + F_{X,Y}(a_1, b_1)$$

Example 2

Lets make a weight matrix used for Gaussian blur. In the weight matrix, each location in the weight matrix will be given a weight based on the probability density of the area covered by that grid square in a Bivariate Normal of independent X and Y, each zero mean with variance σ^2. For this example lets blur using $\sigma = 3$.



Each pixel is given a weight equal to the probability that X and Y are both within the pixel bounds. The center pixel covers the area where $-0.5 \leq x \leq 0.5$ and $-0.5 \leq y \leq 0.5$ What is the weight of the center pixel?

$$
\begin{aligned}
&P(-0.5 < X < 0.5, -0.5 < Y < 0.5)\\
=&P(X < 0.5, Y < 0.5) - P(X < 0.5, Y < -0.5)\\
&- P(X < -0.5, Y < 0.5) + P(X < -0.5, Y < -0.5)\\
=&\phi\left(\frac{0.5}{3}\right) \cdot \phi\left(\frac{0.5}{3}\right) - 2\phi\left(\frac{0.5}{3}\right) \cdot \phi\left(\frac{-0.5}{3}\right)\\
&+ \phi\left(\frac{-0.5}{3}\right) \cdot \phi\left(\frac{-0.5}{3}\right)\\
=&0.5662^2 - 2 \cdot 0.5662 \cdot 0.4338 + 0.4338^2 = 0.206
\end{aligned}
$$