Covariance

- For a single random variable X, the PMF $P_X(x)$ (or the PDF $f_X(x)$) provides a full description while the mean $\mathbb{E}[X]$ and variance $\mathrm{Var}[X]$ provide simple summaries.
- For a pair of random variables X and Y, the joint PMF $P_{X,Y}(x,y)$ (or the joint PDF $f_{X,Y}(x,y)$) provides a full description while the means $\mathbb{E}[X]$, $\mathbb{E}[Y]$ and the variances $\mathrm{Var}[X]$, $\mathrm{Var}[Y]$ summarize how X and Y behave individually.
- Can we summarize the dependencies between X and Y with a single number?
  - → There are many ways to do this.
  - → We will focus on the covariance $\mathrm{Cov}[X,Y]$, which captures the (average) linear relationship between X and Y.

- The covariance $\mathrm{Cov}[X, Y]$ of random variables X and Y is

$$\boxed{\mathrm{Cov}[X,Y] = \mathbb{E}\big[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\big]}$$

→ Sometimes easier to work with by writing

$$\mathrm{Cov}[X,Y] = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)]$$

where $\mu_X = \mathbb{E}[X]$ and $\mu_Y = \mathbb{E}[Y]$.

→ Captures (average) linear relationship between $X - \mathbb{E}[X]$ and $Y - \mathbb{E}[Y]$.

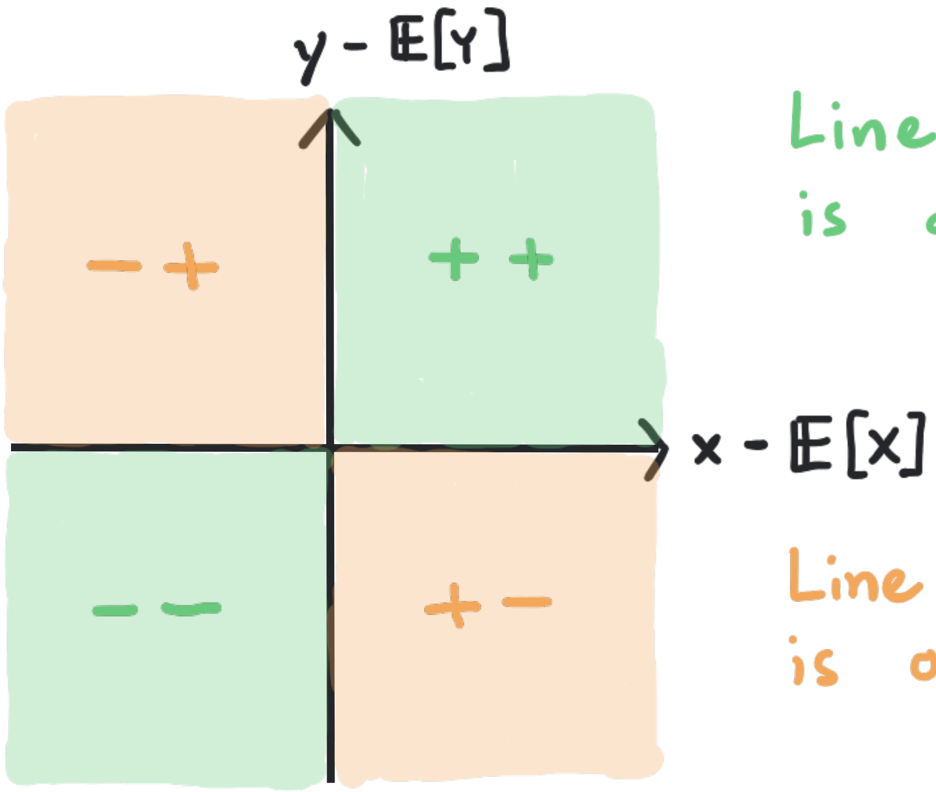
→ Useful Alternate Formula: $\boxed{\mathrm{Cov}[X,Y] = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]}$

Why does this hold?

$$\begin{aligned}
\mathrm{Cov}[X,Y] &= \mathbb{E}[(X-\mu_X)(Y-\mu_Y)] \\
&= \mathbb{E}[XY - X\mu_Y - \mu_X Y + \mu_X\mu_Y] \\
&= \mathbb{E}[XY] - \mathbb{E}[X]\mu_Y - \mu_X\,\mathbb{E}[Y] + \mu_X\mu_Y \\
&= \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]
\end{aligned}$$

Linearity of Expectation →

- **Intuition:** If $\mathrm{Cov}[X,Y] > 0$, then $(X - \mathbb{E}[X])$ and $(Y - \mathbb{E}[Y])$ tend to have the same signs.
  If $\mathrm{Cov}[X,Y] < 0$, then $(X - \mathbb{E}[X])$ and $(Y - \mathbb{E}[Y])$ tend to have opposite signs.
  (This is in a weighted average sense, meaning that larger values of x and y contribute more to the covariance.)



Line with positive slope is a better fit.

Line with negative slope is a better fit.

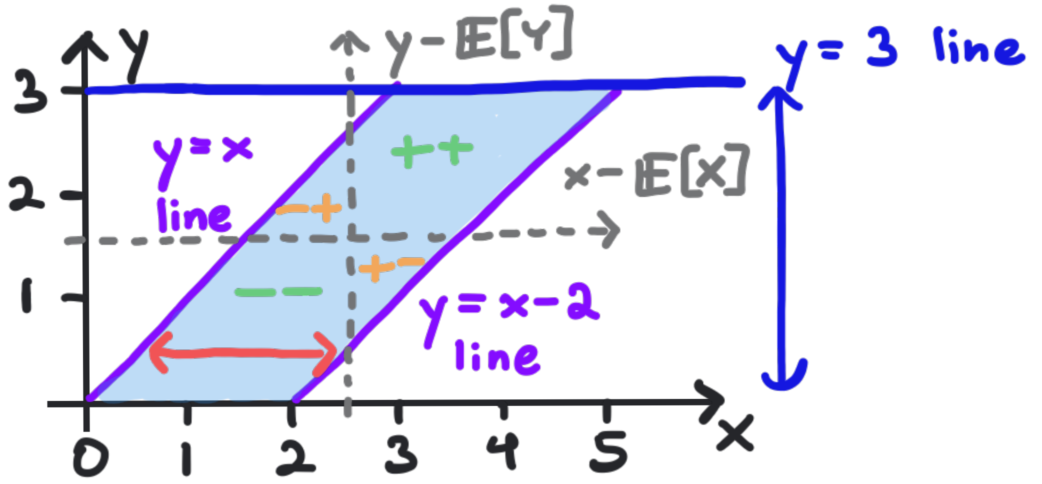
• **Example:** X given $Y=y$ is Uniform$(y, y+2)$. Y is Uniform$(0,3)$. What is the covariance of X and Y?

① Calculate means.

Law of Total Expectation

$$\mathbb{E}[Y] = \frac{0+3}{2} = \frac{3}{2} \qquad \mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}\left[\frac{Y+Y+2}{2}\right] = \mathbb{E}[Y+1] = \frac{5}{2}$$

② Sketch the range.



③ Calculate the covariance.

$$\mathrm{Cov}[X,Y] = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(x-\tfrac{5}{2}\right)\left(y-\tfrac{3}{2}\right) f_{X,Y}(x,y)\,dx\,dy$$

$$= \int_{0}^{3}\int_{y}^{y+2}\left(x-\tfrac{5}{2}\right)\left(y-\tfrac{3}{2}\right)\cdot\tfrac{1}{2}\cdot\tfrac{1}{3}\,dx\,dy$$

$$= \frac{3}{4}$$

$$f_{X,Y}(x,y) = f_{X|Y}(x|y)\,f_Y(y)$$

$$f_Y(y) = \begin{cases} \frac{1}{3} & 0 \le y \le 3 \\ 0 & \text{otherwise} \end{cases}$$

$$f_{X|Y}(x|y) = \begin{cases} \frac{1}{2} & y \le x \le y+2,\ 0 \le y \le 3 \\ 0 & \text{otherwise} \end{cases}$$

→ On average, $X-\mathbb{E}[X]$ and $Y-\mathbb{E}[Y]$ tend to have the same sign: a line with positive slope fits better.

- By the linearity of expectation, $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$.
- What about the variance of the sum?

$$\boxed{\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\,\mathrm{Cov}[X,Y]}$$

→ Why?

$$\begin{aligned}
\mathrm{Var}[X+Y] &= \mathbb{E}\left[(X+Y-\mathbb{E}[X+Y])^2\right] \\
&= \mathbb{E}\left[((X-\mathbb{E}[X]) + (Y-\mathbb{E}[Y]))^2\right] \\
&= \mathbb{E}\left[(X-\mathbb{E}[X])^2 + 2(X-\mathbb{E}[X])(Y-\mathbb{E}[Y]) + (Y-\mathbb{E}[Y])^2\right] \\
&= \mathrm{Var}[X] + 2\,\mathrm{Cov}[X,Y] + \mathrm{Var}[Y]
\end{aligned}$$

Linearity of Expectation

- <u>Variance of Linear Functions:</u>

$$\boxed{\mathrm{Var}[aX+bY+c] = a^2\,\mathrm{Var}[X] + b^2\,\mathrm{Var}[Y] + 2ab\,\mathrm{Cov}[X,Y]}$$

- <u>Covariance of Linear Functions:</u>

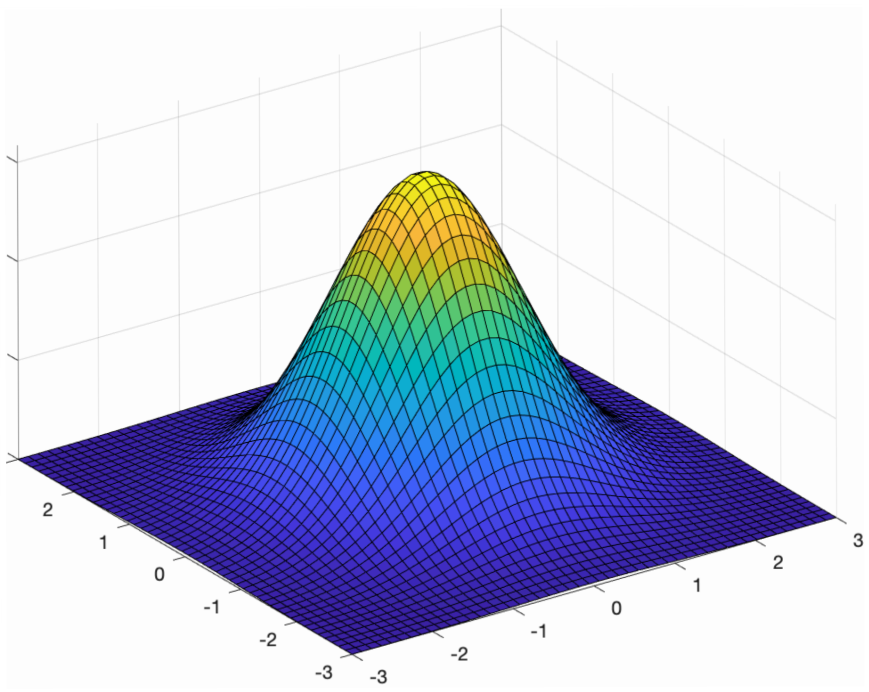
Let $U = aX + bY + c$ and $V = dX + eY + f$. Then,

$$\boxed{\mathrm{Cov}[U,V] = ad\,\mathrm{Var}[X] + be\,\mathrm{Var}[Y] + (ae+bd)\,\mathrm{Cov}[X,Y]}$$

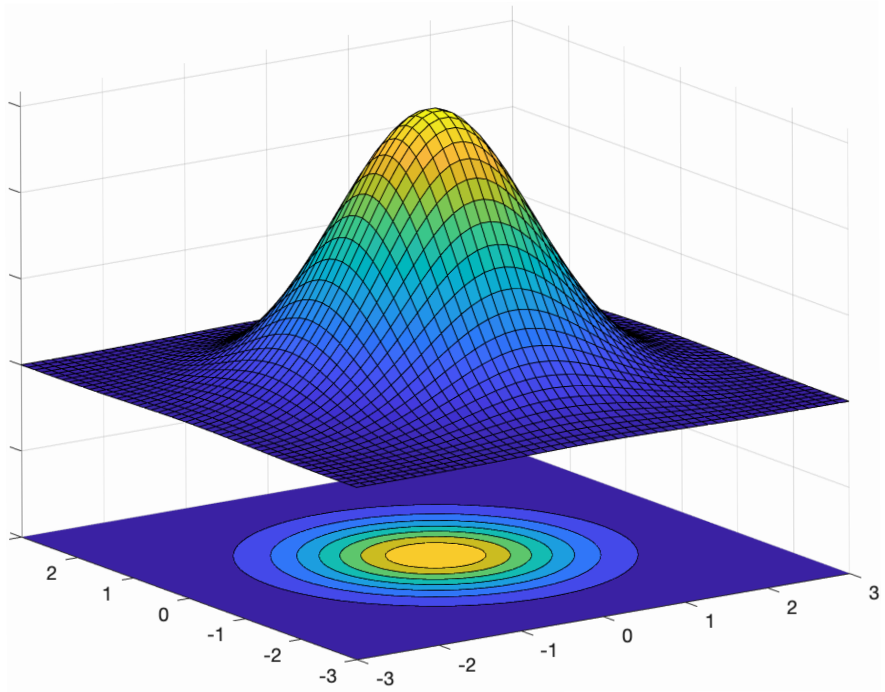
- **Basic Covariance Properties:**
  - → $\mathrm{Cov}[X, Y] = \mathrm{Cov}[Y, X]$
  - → $\mathrm{Cov}[X, X] = \mathrm{Var}[X]$
  - → $\mathrm{Cov}[X, a] = 0$ for any fixed number a.
- X and Y are uncorrelated if $\mathrm{Cov}[X, Y] = 0$.
  - → Independence implies uncorrelatedness.
  - → Uncorrelatedness does not imply independence.
- If X and Y are uncorrelated, then
  - → $\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$
  - → $\mathrm{Var}[aX + bY + c] = a^2\,\mathrm{Var}[X] + b^2\,\mathrm{Var}[Y]$
  - → $\mathrm{Cov}[aX + bY + c, dX + eY + f] = ad\,\mathrm{Var}[X] + be\,\mathrm{Var}[Y]$
  - → $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$

- Visualizing joint PDFs with contour plots:
  - → We can use color to denote the height of a joint PDF.
  - → The contour plot is a top-down view and is often easier to interpret than a surface plot.

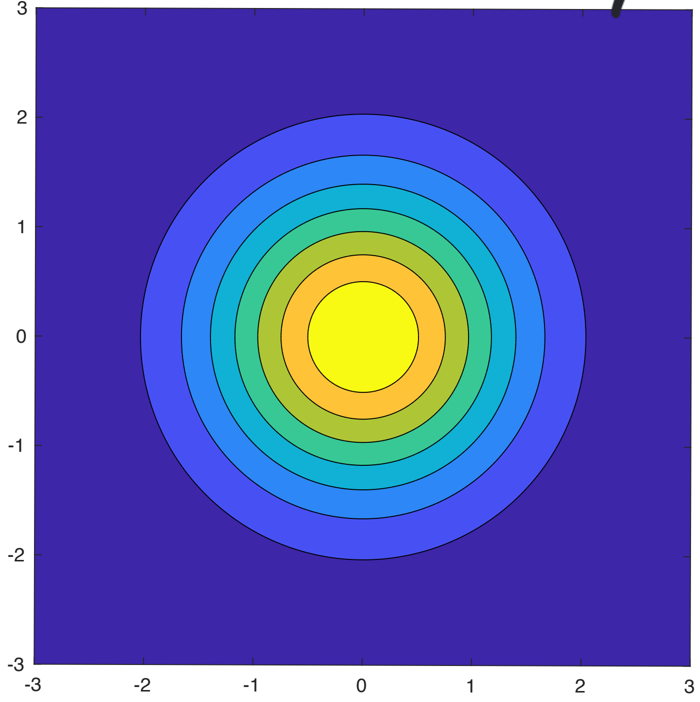
Surface Plot with Color



Contour Plot Underneath

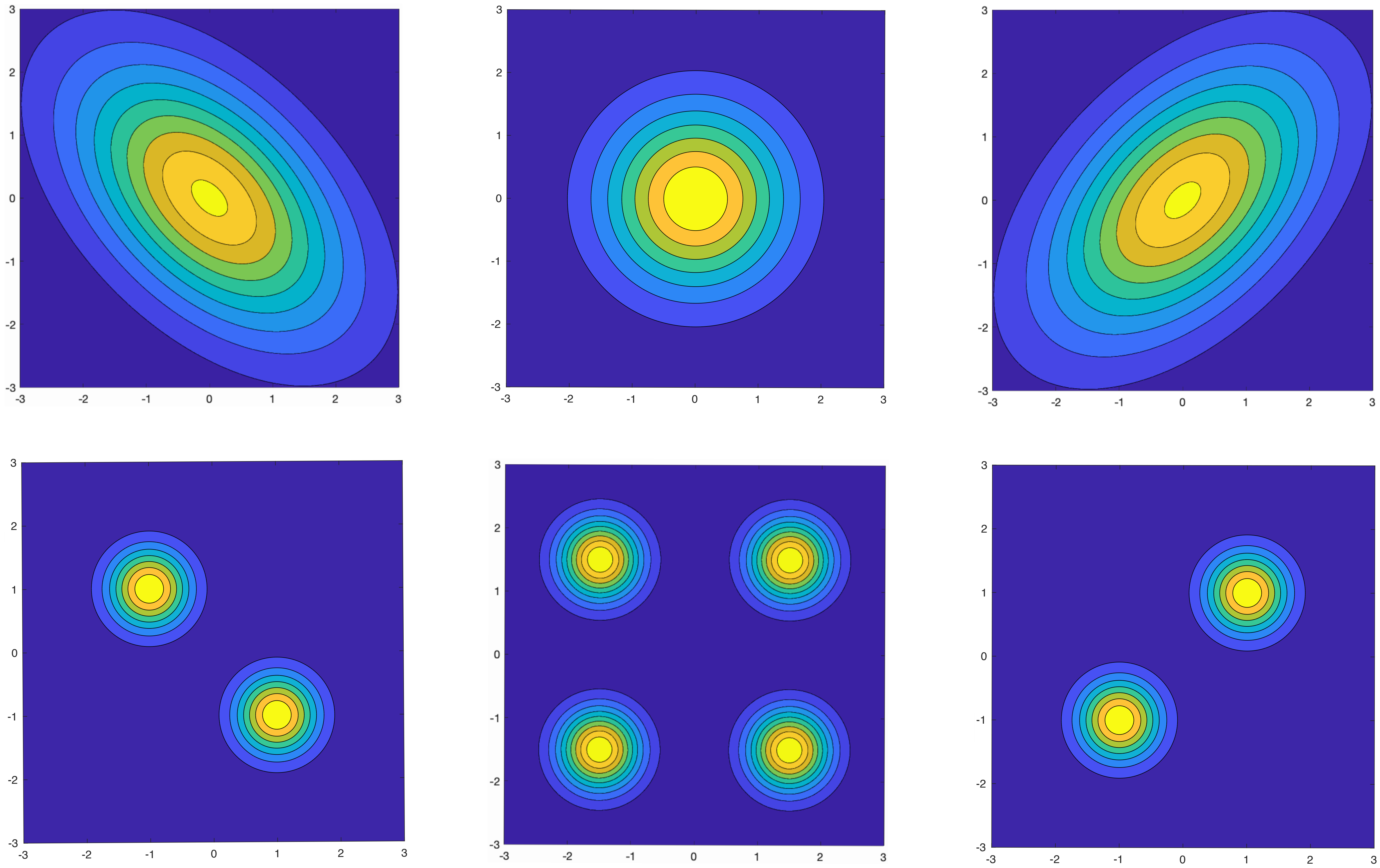


Contour Plot Only



- **Example:**

$Cov[X,Y] = -1$ $\qquad$ $Cov[X,Y] = 0$ $\qquad$ $Cov[X,Y] = +1$



Visualizations are important! Covariance may not tell the full story.