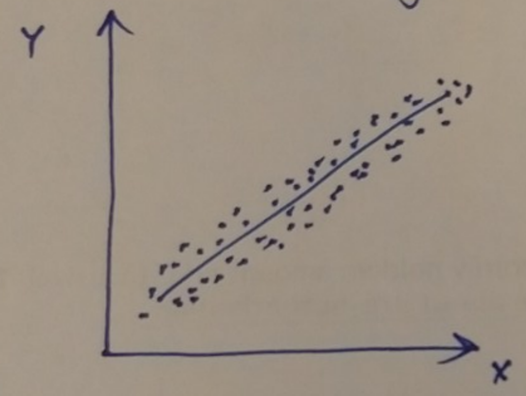
Covariance of Bivariate Gaussian :-



The fig shows that the covariance is positive and strong. The diagonal spread of data is captured by Covariance.

Covariance is given by

$$\sigma(x, y) = E\left[(x - E(x))(y - E(y))\right]$$

Covariance Matrix for Bivariate

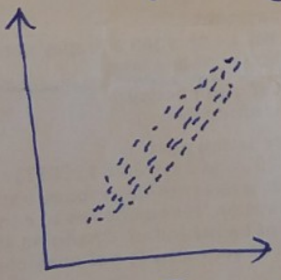
$$\Sigma = \begin{bmatrix} \sigma(x,x) & \sigma(x,y) \\ \sigma(y,x) & \sigma(y,y) \end{bmatrix}$$

If x is positively correlated with y, then it means y is also positively correlated to x.

Thus $\sigma(x,y) = \sigma(y,x)$

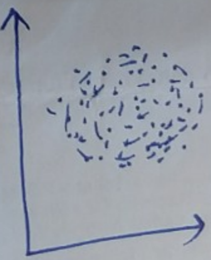
Now from above equ we can see that the matrix is symmetric matrix having variances in diagonal position and off-diagonal contains co-variances.

The following graph depicts the overall shape of the data defining the covariance matrix for Bivariate

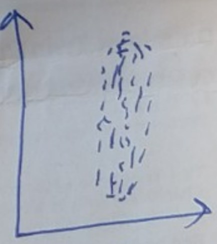


$$\Sigma = \begin{bmatrix} 5 & 4 \\ 4 & 6 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 5 & -4 \\ -4 & 6 \end{bmatrix}$$



$$\Sigma = \begin{bmatrix} \cancel{0}5 & \cancel{5}0 \\ 0 & 1 \end{bmatrix}$$



$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}$$

$X \sim N(\mu_x, \Sigma_x)$ and $Y = Ax + b$ where $b \sim N(0, \Sigma_b)$. Since X and b is from normal distribution Y and $(X,Y)^T$ are also from normal distribution. To find parameters we only need to know its mean and variance.

Now $Y = Ax + b$

mean, $E(Y) = A\,E(X) + E(b) = A\,E(x) = A\mu_x$

Variance, $\Sigma_Y = E[Y - E(Y)]^T (Y - E(Y)] + Var(b)$

$$= E\left[(Ax + b - AE(x) - b)^T (Ax + b - AE(x) - b\right] + \Sigma_b$$

$$= E\left[(AX - AE(X))^T (AX - AE(X)\right] + \Sigma_b$$

$$= A^T E\left[(X - E(X))^T (X - E(X))\right] A + \Sigma b$$

$$= A^T E_X A + \Sigma_b$$

We can say that $Y = Ax + b$ is a normal random ~~variable~~ vector with mean $A\mu_x$ and variance $A^T E_x A + \Sigma_b$

The joint distribution of $(x,Y)^T$ is also normal distribution.

It means, $\mu_{xy} = \begin{pmatrix} \mu_x \\ A\mu_x \end{pmatrix}$

$$Cov(x, Y) = E[XY^T] - E[X]E[Y^T]$$

$$= E\left[X(Ax + b)^T - \mu_x A^T \mu_x^T\right]$$

$$= E[XAx + bx]^T - \mu_x A^T \mu_x^T$$

$$= E[XT^T]A^T - \mu_x A^T \mu_x^T = \Sigma_x A^T$$

Now, $cov(Y,X) = E[X^T Y] - E[X^T] E[Y]$

$$= E[X^T (Ax+b) - \mu_x^T A \mu_x]$$

$$= E[X^T A_x - \mu_x^T A \mu_x]$$

$$= E[X^T x - \mu_x^T \mu_x] A = \Sigma_x A$$

In this derivation we use that the covariance between x and b are zero. Now the variance matrix is,

$$\Sigma_{xy} = \begin{bmatrix} Var(x) & Cov(x,y) \\ Cov(Y,x) & Var(Y) \end{bmatrix} = \begin{bmatrix} \Sigma_x & \Sigma_x A^T \\ A\Sigma_x & A^T \Sigma_x A + \Sigma_b \end{bmatrix}$$