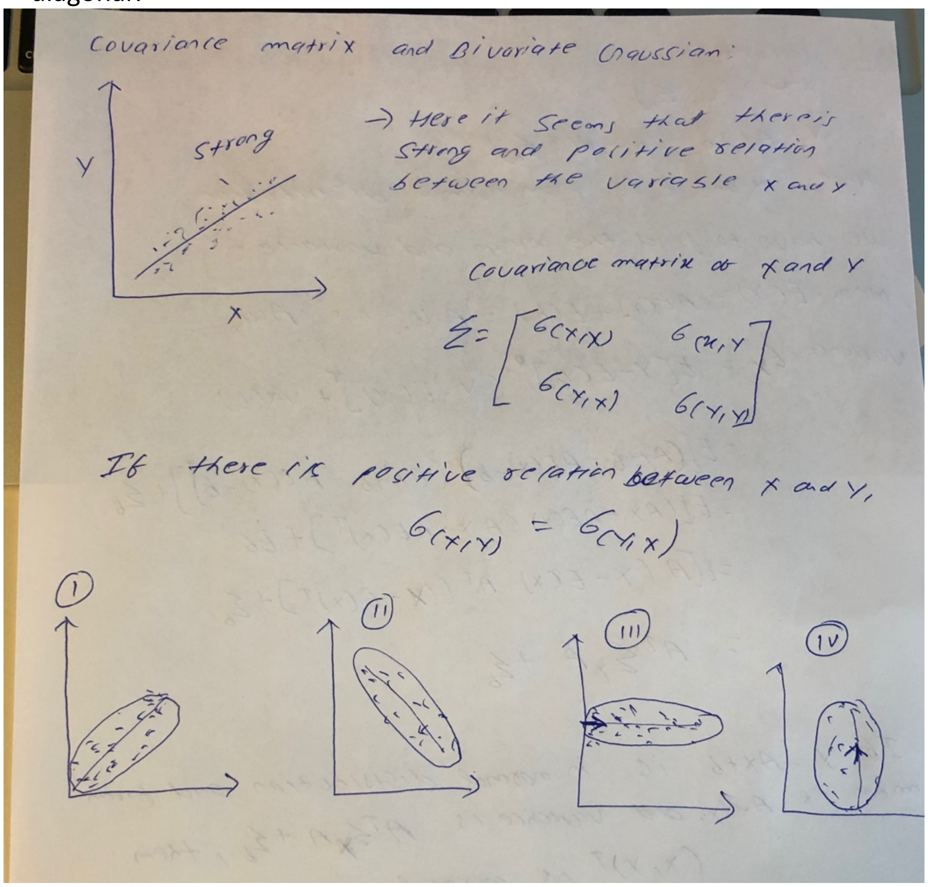
Assignment 2 : Bivariate Gaussian + Linear Transformation of Gaussians

1. Explain in your own words what effect does the choice of Covariance matrix have on the Bivariate Gaussian (compare spherical, elliptical). What does it mean when the covariance matrix is not diagonal?



If the covariance is not diagonal it means there is certain relationship between the variables. The relation can be positive or negative and weak or strong.

2. Do the linear transformation assignment here

② Given

$X \sim N(\mu_x, \Sigma_x)$

$Y = Ax + b$, $b \sim N(0, \Sigma_b)$

Here, x and b is from normal distribution

we have to find the mean and variance

Mean, $E(Y) = AE(x) + E(b) = A \cdot \mu_x + 0 = A\mu_x$

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

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

$$= E[(Ax - AE(x))(Ax - AE(x))^T] + E_b$$

$$= E[A(x - E(x)) A^T ((x - E(x))^T] + \Sigma_b$$

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

If $Y = Ax + b$ is in normal distribution and from mean is $A\mu_x$ and variance is $A^T \Sigma_x A + \Sigma_b$, then $(x, y)^T$ is normal as well.

$$cov(x, y) = E[xy^T] - E[x]E[y^T]$$

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

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

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

now, $cov(Y, x) = E[x^T Y] - E[x^T]E(Y)$

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

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

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

$$= \Sigma_x A.$$

covariance matrix

$$\Sigma_{xy} = \begin{bmatrix} var(x) & cov(x, y) \\ cov(y, x) & var(y) \end{bmatrix}$$

$$= \begin{bmatrix} \Sigma_x & \Sigma_x A^T \\ A\Sigma_x & A^T \Sigma_x A + \Sigma_b \end{bmatrix}$$

3. What is the meaning of Mahalanobis distance? What is the relation of this to the eigenvalues of the Covariance matrix? Draw a sketch either in Python or by hand for the Bivariate case (K=2)

   Mahalanobis distance is distance between any points in the multivariate space. It remove the error given by the Euclidean distance. There few features of the Mahalanobis distance like it removes the colinearinity of the variables, makes the variances equals to 1 and makes the variables independent. The distance is given by:
   D=(x-m)^T(C^-1)(x-m) , where x is data point, m is mean and c^-1 is the inverse covariance matrix.

If the covariances are zero, then the eigenvalues are equal to the variances: If the covariance matrix not diagonal, the eigenvalues represent the variance along the principal components, whereas the covariance matrix still operates along the axes: