Multivariate Probability Density fn:

Here X be a feature vector.

$X = \langle x_1, x_2, \ldots, x_d \rangle$

$$P(X) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left[-\frac{1}{2} (X-M)^t \Sigma^{-1} (X-M)\right]$$

$M = E(X)$; x – d-dimensional ;

$$\text{Mean vector} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$$

$$\boxed{M = \int_{-\infty}^{\infty} X \; P(X) \, dx}$$

Σ = covariance Matrix.

$$= E\left[(X-M)(X-M)^t\right] = (d \times 1) \; (1 \times d) = (d \times d)$$

matrix

$$\boxed{\Sigma = \int_{-\infty}^{\infty} (X-M)(X-M)^t \; P(X) \, dx.}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} \Rightarrow \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$C = \begin{bmatrix} cov(x,x) & cov(x,y) & cov(x,z) \\ cov(y,x) & cov(y,y) & cov(y,z) \\ cov(z,x) & cov(z,y) & cov(z,z) \end{bmatrix}$$

⇒ Diagonal is the variances of x, y and z.

⇒ $cov(x,y) = cov(y,x)$ hence matrix is symmetrical about the diagonal

⇒ N- dimensional data will result in NXN matrix

$$C = \begin{array}{c|ccc} & 1 & 2 & d \\ 1 & E[(x_1-M_1)(x_1-M_1)] & E[(x_1-M_1)(x_2-M_2)] & E[(x_1-M_1)(x_d-M_d)] \\ 2 & E[(x_2-M_2)(x_1-M_1)] & E[(x_2-M_2)(x_2-M_2)] & E[(x_2-M_2)(x_d-M_d)] \\ \vdots & \vdots & & \\ d & E[(x_d-M_d)(x_1-M_1)] & E[(x_d-M_d)(x_2-M_2)] & E[(x_d-M_d)(x_d-M_d)] \end{array}$$

The definition above is equivalent to the matrix equality.

$$\boxed{C = \Sigma = E[(X-M)\cdot(X-M)^t]}$$

(What is expected value of individual component)

Expected value of ith component:

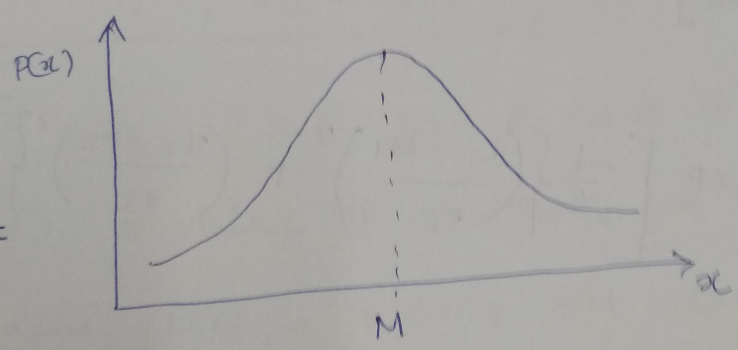
$$\mu_i = E[x_i]$$

$$\sigma_{ij} = E[(x_i - \mu_i)(x_j - \mu_j)]$$

$$i = j; \quad \sigma_{ii} = E[(x_i - \mu_i)^2]$$

$$\sigma_i^2 = E[(x_i - \mu_i)^2]$$

Let us consider the case of single variable normal density [uni-variate normal density]



Before multi-variate density case, let us see what will happen in 2-D case?

What does this multivariate normal density actually tell us? (or) What is the interpretation of multi-variate density?

Bi-variate normal density:

any two variables x_1 and x_2

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$P(X) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left[-\frac{1}{2} (X - \mu)^t \Sigma^{-1} (X - \mu) \right]$$

can be simplified to

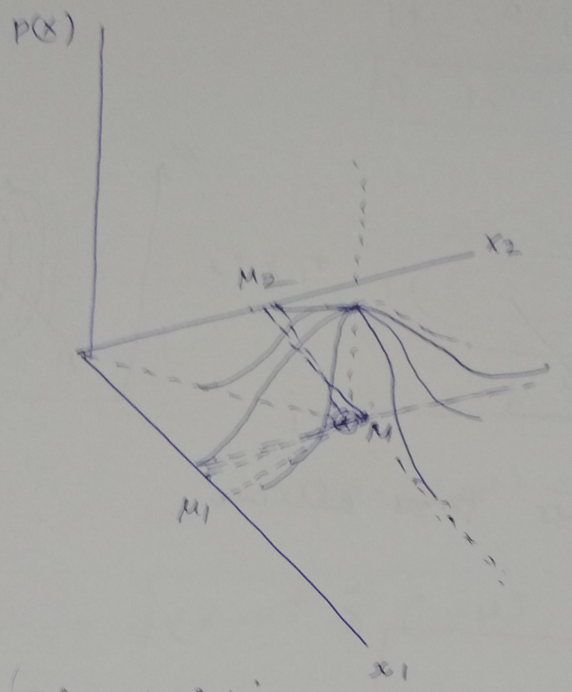
$\boxed{d=2}$; Assume x_1 and x_2 are statistically independent

σ_{12} and σ_{21} are '0'.

$$\underset{\text{Mean vector}}{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \text{ and } \Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

$$P(X) = \frac{1}{2\pi |\Sigma|^{1/2}} \exp\left[-\frac{1}{2} \left\{ \left(\frac{x_1 - \mu_1}{\sigma_1} \right)^2 + \left(\frac{x_2 - \mu_2}{\sigma_2} \right)^2 \right\} \right]$$

2-dimensional Feature Space:



$$a^m \cdot a^n = a^{m+n}$$

$\left\{ \begin{array}{l} \text{addition of two} \\ \text{exponential term is} \\ \text{equivalent product} \end{array} \right.$

$\boxed{\text{Case 1}}$ $\sigma_{ij} = 0 \; ; \; i \neq j$

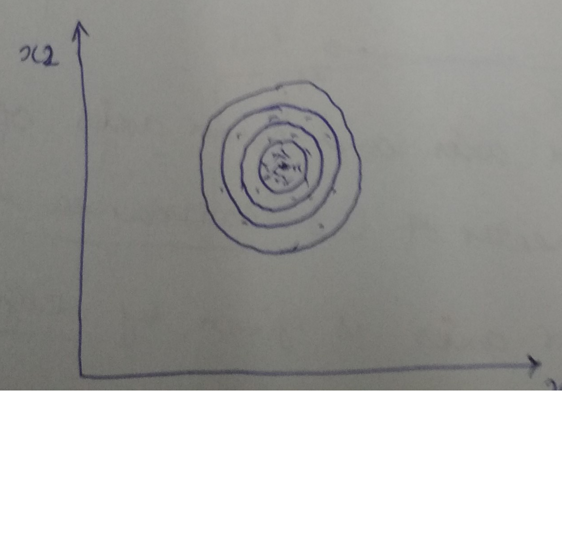
$$\boxed{\sigma_1^2 = \sigma_2^2 \; ; \; \sigma_{12} = \sigma_{21} = 0}$$

Trace, the loci of points of constant density for

(i.e) $\forall$ value of x; for which $P(x)$ is constant

Those loci is nothing but circle.

Footprint on x_1 and x_2 plane:

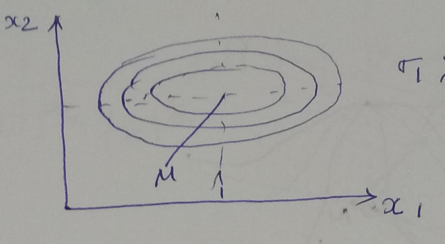


What happens if the variance are different?

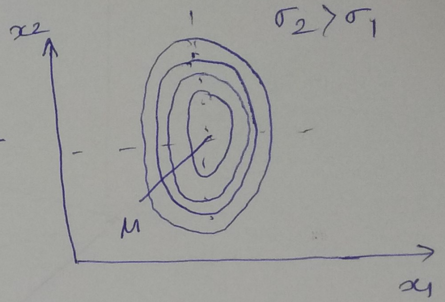
Case 2: $\sigma_1^2 \neq \sigma_2^2$

$\sigma_{ij} = 0;\ i \neq j$

$\boxed{\sigma_{12} = \sigma_{21} = 0}$



$\sigma_1 > \sigma_2$



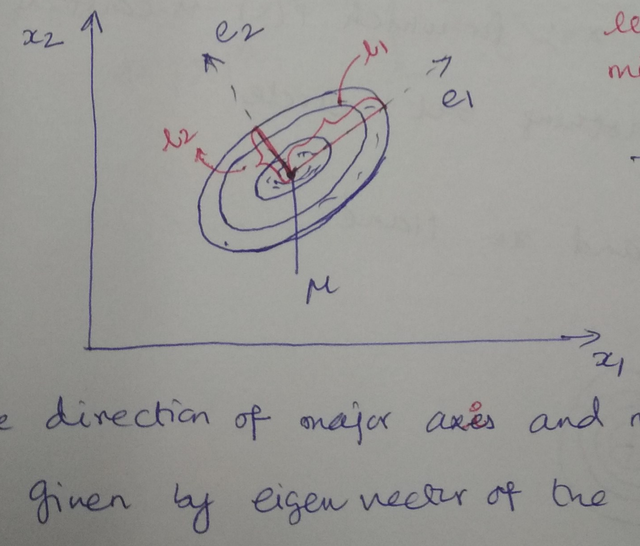
$\sigma_2 > \sigma_1$

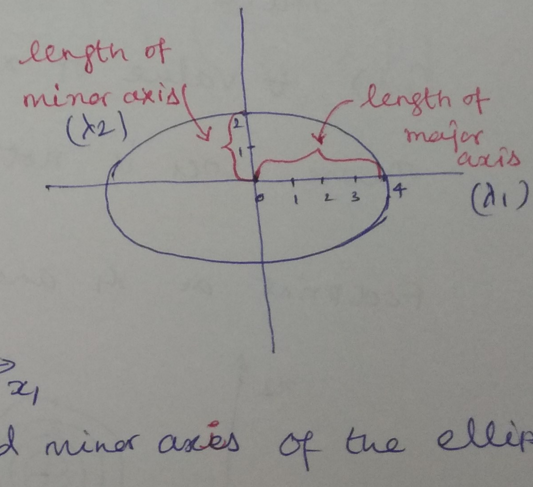
The loci of points forms ellipse.

Case 3:

$\sigma_{ij} \neq 0 \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$

x_1 and x_2 are not statistically independent.





The direction of major axes and minor axes of the ellips will be given by eigen vector of the covariance matrix Σ.

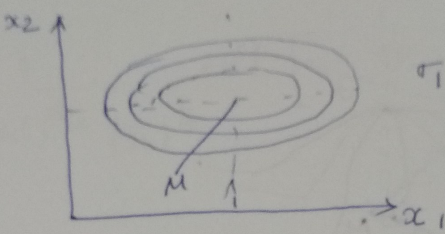
The length of major & minor axes is given by eigen values.

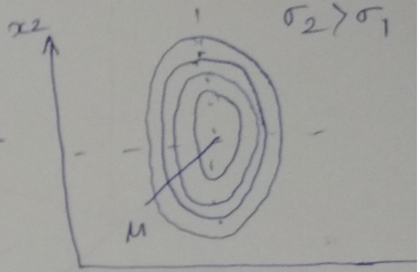
what happens if the variance are different?

Case 2: $\sigma_1^2 \neq \sigma_2^2$

$\sigma_{ij} = 0;\ i \neq j$

$\boxed{\sigma_{12} = \sigma_{21} = 0}$



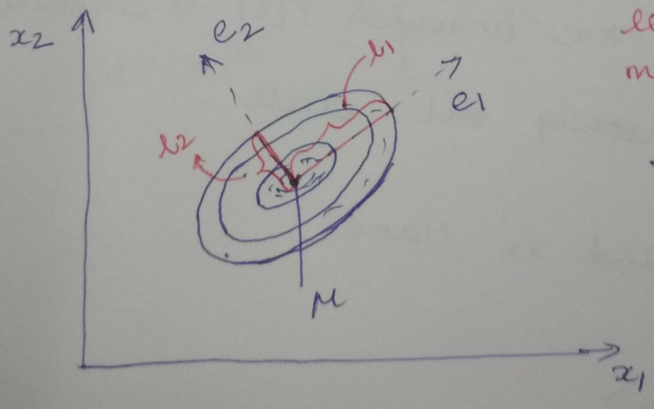


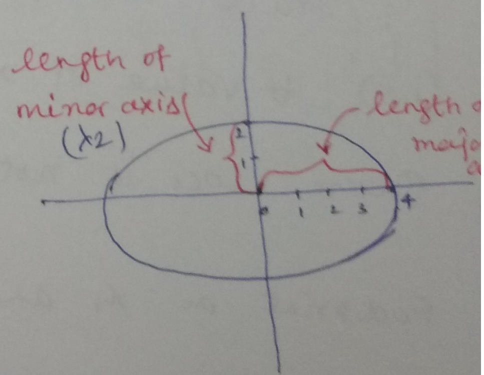
The loci of points forms ellipse.

Case 3:

$\sigma_{ij} \neq 0 \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$

x_1 and x_2 are not statistically independent.





The direction of major axes and minor axes of the e
will be given by eigen vector of the covariance matrix. Σ
The length of major & minor axes is given by eigen values

Coming to Multivariate Normal density:

The surface forms an 'ellipsoid' in multivariate normal density

$$(X - M)^t \Sigma^{-1} (X - M) \rightarrow$$ Quadratic form of Loci of Points of constant density.

$$r^2 = (X - M)^t \Sigma^{-1} (X - M) \Rightarrow$$ Mahalanobis distance.

$$r = \sqrt{(X - M)^t \Sigma^{-1} (X - M)}$$

Note:

If the covariance matrix is the identity matrix the, Mahalonobis distance reduces to the Euclidean distance

$$(X - M)^t \cdot (X - M) = (X - M)^2 \text{ d-dimension}$$

$$\ell^2 = \sum_{i=1}^{d} (x_i - M_i)^2$$

$$\ell = \sqrt{\sum_{i=1}^{d} (x_i - M_i)^2}$$

If the covariance matrix is diagonal, then the resulting distance measure is called a <u>standardized Euclidean distance</u>.