Theorem 4.4 (Cramer-Wold device)

The distribution of $\underline{X}$ is completely determined by the set of all (one dimensional) distribution of $\underline{a}^T\underline{x}$

Proof. $M_{\underline{a}^T\underline{x}}(s) = E[e^{s\underline{a}^T\underline{x}}]$

$$M_{\underline{x}}(\underline{t}) = E[e^{\underline{t}^T\underline{x}}] = M_{\underline{t}^T\underline{x}}(1)$$

which shows that the momentgenerating function for $\underline{X}$ for a given $\underline{t}$ can be found from the univariate mgf of $\underline{t}^T\underline{x}$.

The multivariate normal distribution

Definition. A random vector $\underline{X}_{p\times 1}$ is multivariate normal distributed if $\underline{a}^T\underline{X}$ is univariate normal distributed $\forall \underline{a} \in R^p$ (p is the dimension of $\underline{X}$).

Theorem 5.2. Let $\underline{X}$ be multivariate normal. Then all linear transformations of $\underline{X}$ followed by a translation i.e $\underline{A}\,\underline{X} + \underline{c}$ is multivariate normal distributed.

Proof. Let $\underline{Y} = \underline{A}\,\underline{X} + \underline{c}$. $\underbrace{\underline{a}^T\underline{Y}}_{\forall \underline{a} \in R^q \text{ (dimension of } \underline{Y})} = \underline{a}^T(\underline{A}\,\underline{X} + \underline{c}) = \underline{a}^T\underline{A}\,\underline{X} + \underline{a}^T\underline{c}$

$= (\underline{A}^T\underline{a})^T\underline{X} + \underline{a}^T\underline{c}$. $\quad \underbrace{(\underline{A}^T\underline{a})}_{\underline{a}^*}{}^T\underline{X} = \underbrace{\underline{a}^{*T}\underline{X}}_{\text{univariate.}} + \underline{a}^T\underline{c}$. $\quad \underline{a}^T\underline{c}$ is a number.

This implies that $\underline{a}^T(\underline{A}\,\underline{X} + \underline{c})$ is univariate normal distributed for a $\underline{a}$ and therefore $\underline{Y} = \underline{A}\,\underline{X} + \underline{c}$ is multivariate normal distributed.

$E[\underline{Y}] = \underline{A}\,E[\underline{X}] + \underline{c} = \underline{A}\,\underline{\mu} + \underline{c}$ and $Cov(\underline{Y}) = \underline{A}\,\underline{\Sigma}\,\underline{A}^T$.

Theorem. $\underline{X} \sim N(\underline{\mu}, \underline{\Sigma})$ with $\underline{\Sigma}$ positive definit.

Then $f(\underline{x}) = \dfrac{1}{(2\pi)^{\frac{p}{2}}\,|\underline{\Sigma}|^{\frac{1}{2}}}\, e^{-\frac{1}{2}(\underline{x}-\underline{\mu})^T\underline{\Sigma}^{-1}(\underline{x}-\underline{\mu})}$

Proof. Let $\underline{Z} = [Z_1, \dots, Z_p]^T \sim N(\underline{0}, \underline{I})$

$\Rightarrow f_{\underline{Z}}(\underline{z}) = \dfrac{1}{(2\pi)^{\frac{p}{2}}}\, e^{-\frac{1}{2}\underline{z}^T\underline{z}} \qquad \left(f_{Z_i}(z) = \dfrac{1}{\sqrt{2\pi}}\, e^{-\frac{z^2}{2}} \right)$

Define $\underline{X} = \underline{\Sigma}^{\frac{1}{2}}\underline{Z} + \underline{\mu}$ i.e $\underline{Z} = \underline{\Sigma}^{-\frac{1}{2}}(\underline{X} - \underline{\mu})$ (Mahalanobis transformation)

Then $E[\underline{X}] = \underline{\Sigma}^{\frac{1}{2}}\underbrace{E[\underline{Z}]}_{\underline{0}} + \underline{\mu} = \underline{\mu}$, $Cov(\underline{X}) = \underline{\Sigma}^{\frac{1}{2}}\underline{I}\,\underline{\Sigma}^{\frac{1}{2}} = \underline{\Sigma}$

Hence $\underline{X} \sim N(\underline{\mu}, \underline{\Sigma})$ and

$$f_{\underline{X}}(\underline{x}) = f_{\underline{Z}}\left(\underline{\Sigma}^{-\frac{1}{2}}(\underline{X} - \underline{\mu})\right)\cdot|\underline{\Sigma}|^{-\frac{1}{2}} = 2\pi^{-\frac{p}{2}}|\underline{\Sigma}|^{-\frac{1}{2}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu})^T\underline{\Sigma}^{-1}(\underline{x}-\underline{\mu})}$$

Result 6. All vectors of components from $\underline{X}$ (multivariate normal) are also multivariate normally distributed.

Proof. Choose an appropriate $\underline{A}$ with 0 and 1 as elements and let $\underline{c} = \underline{0}$. For instance for $\underline{X} = [x_1, x_2, \dots, x_5]^T$

$$\begin{bmatrix} X_1 \\ X_2 \\ X_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \end{bmatrix}$$

and $\underline{Y} = \underline{A}\,\underline{X}$ is multivariate normal according to theorem 5.2

$$\underline{Y} = \underline{A}\,\underline{X}.$$

Quadratic forms as distance measures

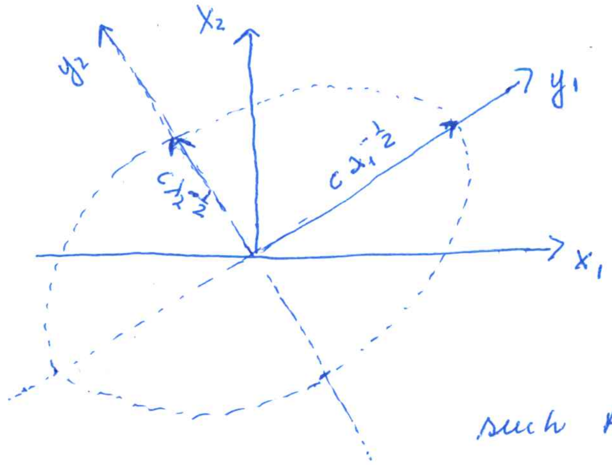
Let $A_{k\times k}$ be symmetric and positive definite i.e $\underline{x}^T\underline{A}\,\underline{x} > 0 \quad \forall \underline{x} \neq \underline{0}$. Every point with a distance c from origo satisfies $\underline{x}^T\underline{A}\,\underline{x} = c^2$. In particular for $\underline{A}_{2\times 2}$ we get $\underline{x}^T\underline{A}\,\underline{x} = c^2 \iff [x_1, x_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{21}x_2^2 = c^2$

Also $\underline{A} = \underline{P}\,\underline{\Lambda}\,\underline{P}^T$ where $P = [\underline{e}_1, \underline{e}_2]$, the matrix of normalized eigenvectors and $\underline{\Lambda} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ (eigenvalues)

Hence $\underline{A} = \lambda_1 \underline{e}_1 \underline{e}_1^T + \lambda_2 \underline{e}_2 \underline{e}_2^T$

and $\underline{x}^T \underline{A}\, \underline{x} = \lambda_1 \underbrace{\underline{x}^T \underline{e}_1}_{y_1} \underline{e}_1^T \underline{x} + \lambda_2 \underline{x}^T \underline{e}_2 \underline{e}_2^T \underline{x} = \lambda_1 y_1^2 + \lambda_2 y_2^2$

and $\underline{x}^T \underline{A}\, \underline{x} = c^2$ is the equation for an ellipse in y_1 and y_2



$\underline{x} = c\lambda_1^{-\frac{1}{2}} \underline{e}_1 \Rightarrow \begin{cases} y_1 = c\lambda_1^{-\frac{1}{2}} \\ y_2 = 0 \end{cases}$

$\underline{x} = c\lambda_2^{-\frac{1}{2}} \underline{e}_2 \Rightarrow \begin{cases} y_1 = 0 \\ y_2 = c\lambda_2^{-\frac{1}{2}}. \end{cases}$

are points on the ellipse $\lambda_1 y_1^2 + \lambda_2 y_2^2 = c^2$ such that the length of the half axes are $c\lambda_1^{-\frac{1}{2}}$ and $c\lambda_2^{-\frac{1}{2}}$. In general $(\underline{x}-\underline{\mu})^T \underline{A} (\underline{x}-\underline{\mu})$ gives the same squared distance, c^2, for all points that lie on an ellipsoid with half axes given by $c\lambda_i^{-\frac{1}{2}} \underline{e}_i$.

The multivariate normal density is constant on surfaces where $(\underline{x}-\underline{\mu})^T \underline{\Sigma}^{-1} (\underline{x}-\underline{\mu})$ is constant i.e for all $\underline{x}$ such that $(\underline{x}-\underline{\mu})^T \underline{\Sigma}^{-1} (\underline{x}-\underline{\mu}) = c^2$.

<u>Theorem</u> (corolar 5.2 in HS). Assume $\underline{X} \sim \underline{N}(\underline{\mu}, \underline{\Sigma})$.

Then $\underline{A}\underline{X}$ and $\underline{B}\underline{X}$ are independent if and only if $\underline{A}\underline{\Sigma}\underline{B}^T = \underline{0}$

<u>Note</u> $\underline{A}\underline{\Sigma}\underline{B}^T = \mathrm{cov}(\underline{A}\underline{X}, \underline{B}\underline{X})$

$\Rightarrow$ $\underline{A}\underline{X}$ and $\underline{B}\underline{X}$ independent $\Rightarrow \mathrm{Cov}(\underline{A}\underline{X}, \underline{B}\underline{X}) = \underline{A}\underline{\Sigma}\underline{B}^T = \underline{0}$

$\Leftarrow$ $\begin{bmatrix} \underline{A} \\ \underline{B} \end{bmatrix} \underline{X} \sim N\left(\begin{bmatrix} \underline{A} \\ \underline{B} \end{bmatrix} \underline{\mu}, \begin{bmatrix} \underline{A} \\ \underline{B} \end{bmatrix} \underline{\Sigma} \begin{bmatrix} \underline{A}^T & \underline{B}^T \end{bmatrix} \right)$, $\begin{bmatrix} \underline{A} \\ \underline{B} \end{bmatrix} \underline{\Sigma} \begin{bmatrix} \underline{A}^T & \underline{B}^T \end{bmatrix} = \begin{bmatrix} \underline{A}\underline{\Sigma}\underline{A}^T & \underline{A}\underline{\Sigma}\underline{B}^T \\ \underline{B}\underline{\Sigma}\underline{A}^T & \underline{B}\underline{\Sigma}\underline{B}^T \end{bmatrix}$