THE MULTIVARIATE GAUSSIAN DISTRIBUTION.

Density

Let $\underline{\mu} \in \mathbb{R}^N$ and $R \in \mathbb{R}^{N \times N}$ be symmetric and positive definite. A random variable $\underline{X}$ has a <u>multivariate</u> <u>Gaussian</u> <u>distribution</u> if its density is

$$f(\underline{x}) = \frac{1}{(2\pi)^{\frac{N}{2}} \, |R|^{\frac{1}{2}}} \, e^{-\frac{1}{2}(\underline{x}-\underline{\mu})^T R^{-1} (\underline{x}-\mu)}$$

Mean and covariance

$$\underline{\mu} = E[\underline{X}]$$

$$R = E\left[(\underline{X}-\underline{\mu})(\underline{X}-\underline{\mu})^T\right]$$

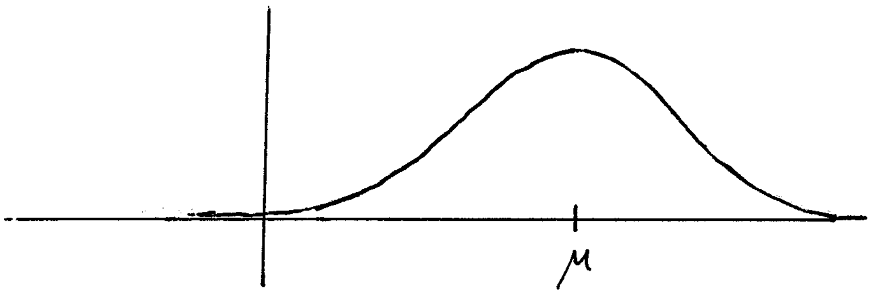
Notation

$$\underline{X} \sim N(\underline{\mu}, R)$$

Conceptualization

In 1-d, $R = [\sigma^2]$ (1×1) and

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$



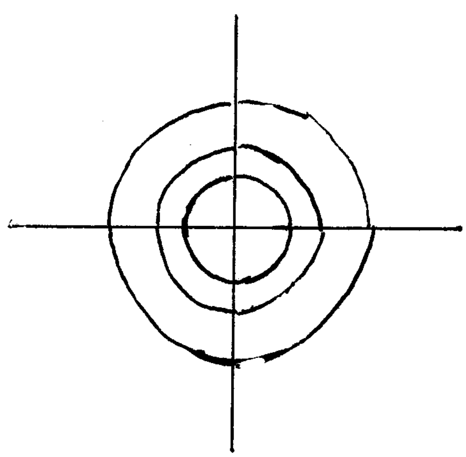
bell curve

In 2-d, let's consider 3 cases:

<u>Case 1</u>: $R = \sigma^2 I_{2\times 2} = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$

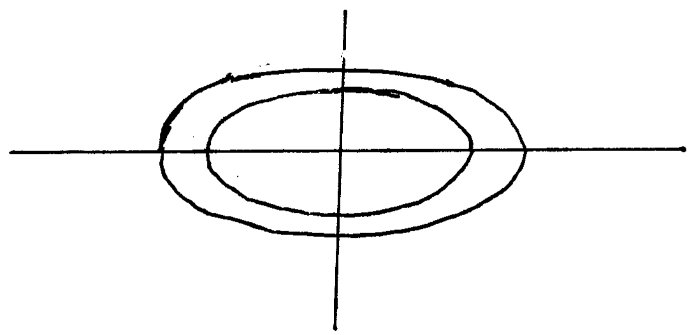
Then a contour of the density is a <u>circle</u>:

$$f(\underline{x}) \equiv \gamma \iff \|\underline{x} - \underline{\mu}\|^2 \equiv \gamma'$$



Case 2 : $R = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$, say $\sigma_1 > \sigma_2$

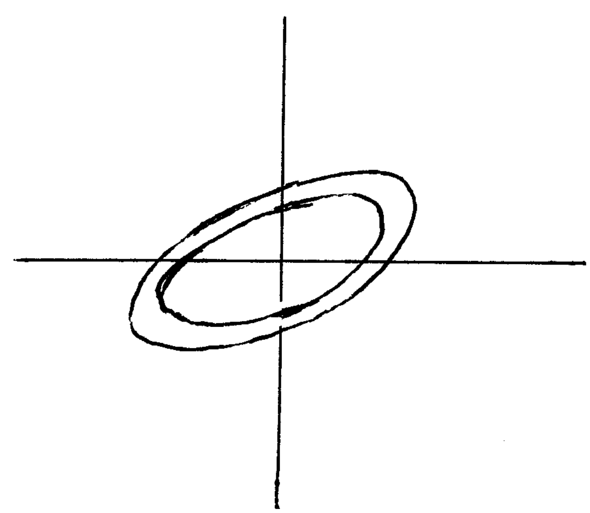
Then the density contours are ellipses whose axes align with the standard basis.



To see this, observe

$$f(\underline{x}) \equiv \gamma \iff \frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2} \equiv \gamma'$$

Case 3 : R is arbitrary. Then density contours are ellipses with arbitrary orientation.



To see this, write

$$R = U \Lambda U^T$$

Then

$$(\underline{x} - \underline{\mu})^T R^{-1} (\underline{x} - \underline{\mu})$$

$$= (\underline{x} - \underline{\mu})^T U \Lambda^{-1} U^T (\underline{x} - \underline{\mu})$$

$$= (\underline{x}' - \underline{\mu}')^T \Lambda^{-1} (\underline{x}' - \underline{\mu}')$$

$$\left[\text{where } \underline{x}' = U^T \underline{x}, \ \underline{\mu}' = U^T \underline{\mu}\right]$$

$$= \frac{(x_1' - \mu_1')^2}{\lambda_1} + \frac{(x_2' - \mu_2')^2}{\lambda_2}$$

which defines an ellipse in the rotated coordinate system.

More generally, the MVG distribution

- is symmetric with respect to its mean
- is unimodal
- has ellipsoidal contours: axes $\Leftrightarrow$ eigenvectors of R and axis lengths $\Leftrightarrow$ eigenvalues of R

Importance

The MVG model is the most important and widely employed model in statistical signal processing.

Some reasons for this include:

- Tractability
- Estimators and detectors with intuitive forms and properties
- Justification in terms of the central limit theorem.

CLT : If $\underline{X} = \frac{1}{n} \sum_{i=1}^{n} \underline{Y}_i$, then

rough paraphrase

$$\underline{X} \to \mathcal{N}(\underline{\mu}, R) \quad \text{as} \quad n \to \infty$$

for some $\underline{\mu}$, R, regardless of the distribution of $\underline{Y}$.

Example) In communication systems, electronic noise is due to the aggregate effect of huge numbers of charge carriers undergoing random motion.

Characteristic function

The characteristic function of an N dimensional random variable $\underline{X}$ is defined to be

$$\Phi(\underline{\omega}) = E[e^{-j\underline{\omega}^T \underline{X}}]$$

$$= \int e^{-j\underline{\omega}^T \underline{x}} f(\underline{x})\, d\underline{x}$$

The char. fun. is an N-dim Fourier transform of the density of $\underline{X}$. Thus it uniquely characterizes the random variable. The density may be recovered from Φ by taking the inverse Fourier transform.

For the MVG, $\underline{X} \sim N(\underline{\mu}, R)$ we have

$$\Phi(\underline{\omega}) = E[e^{-j\underline{\omega}^T\underline{X}}]$$

$$= \int e^{-j\underline{\omega}^T\underline{x}} f(\underline{x})\, d\underline{x}$$

$$= \int (2\pi)^{-\frac{N}{2}} |R|^{-\frac{1}{2}} \exp\left\{-j\underline{\omega}^T\underline{x} - \frac{1}{2}(\underline{x}-\underline{\mu})^T R^{-1}(\underline{x}-\underline{\mu})\right\} d\underline{x}$$

$$\boxed{\text{complete the square}} \longrightarrow \quad = e^{-j\underline{\omega}^T\underline{\mu} - \frac{1}{2}\underline{\omega}^T R\underline{\omega}} \times$$

$$\underbrace{\int (2\pi)^{-\frac{N}{2}} |R|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\underline{x}-\underline{\mu}+jR\underline{\omega})^T R^{-1}(\underline{x}-\underline{\mu}+jR\underline{\omega})\right\} d\underline{x}}_{\text{Gaussian density} \to 1}$$

$$= e^{-j\underline{\omega}^T\underline{\mu} - \frac{1}{2}\underline{\omega}^T R\underline{\omega}}$$

Linear Transformations

Proposition If $\underline{X} \sim \mathcal{N}(\underline{\mu}, R)$ is N-dim, $A \in \mathbb{R}^{M \times N}$, and $\underline{Y} = A\underline{X}$, then

$$\underline{Y} \sim \mathcal{N}(A\underline{\mu}, ARA^T)$$

Proof

$$\begin{aligned}
\Phi_y(\underline{\omega}) &= E\left[e^{-j\underline{\omega}^T \underline{Y}}\right] \\
&= E\left[e^{-j\underline{\omega}^T A\underline{X}}\right] \\
&= E\left[e^{-j(A^T\underline{\omega})^T \underline{X}}\right] \\
&= \Phi_X(A^T\underline{\omega}) \\
&= e^{-j\underline{\omega}^T A\underline{\mu} - \frac{1}{2}\underline{\omega}^T ARA^T \underline{\omega}}
\end{aligned}$$

$$\Rightarrow \underline{Y} \sim \mathcal{N}(A\underline{\mu}, ARA^T)$$

since the characteristic function uniquely characterizes a distribution.

Marginals

Proposition | Let $\underline{X} = \begin{bmatrix} \underline{X}_1 \\ \underline{X}_2 \end{bmatrix}$, $\underline{\mu} = \begin{bmatrix} \underline{\mu}_1 \\ \underline{\mu}_2 \end{bmatrix}$, $R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$.

If $\underline{X} \sim N(\underline{\mu}, R)$, then

$$\underline{X}_1 \sim N(\underline{\mu}_1, R_{11}).$$

Exercise | Prove this.

Solution | Write $\underline{X}_1 = A\underline{X}$ where

$$A = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \quad \begin{matrix} \\ 0 \\ \\ \end{matrix} \qquad (p \times N)$$

assuming $\underline{X} \in \mathbb{R}^N$, $\underline{X}_1 \in \mathbb{R}^p$.

Then

$$A\underline{\mu} = \underline{\mu}_1$$

$$ARA^T = R_{11}$$

Now apply the previous result.

Conditioning

Proposition | Let $\underline{X} = \begin{bmatrix} \underline{X}_1 \\ \underline{X}_2 \end{bmatrix}$, $\underline{\mu} = \begin{bmatrix} \underline{\mu}_1 \\ \underline{\mu}_2 \end{bmatrix}$, $R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$.

If $\underline{X} \sim N(\underline{\mu}, R)$, then

$$\underline{X}_2 \mid \underline{X}_1 = \underline{x}_1 \sim N(\tilde{\underline{\mu}}, \tilde{R})$$

where

$$\tilde{\underline{\mu}} = \underline{\mu}_2 + R_{21} R_{11}^{-1} (\underline{x}_1 - \underline{\mu}_1)$$

$$\tilde{R} = R_{22} - R_{21} R_{11}^{-1} R_{12}$$

Proof Write out $f(\underline{x}_2 | \underline{x}_1) = \dfrac{f(\underline{x})}{f(\underline{x}_1)}$

and simplify. See Kay (Vol. I) or Moon and Stirling for details.