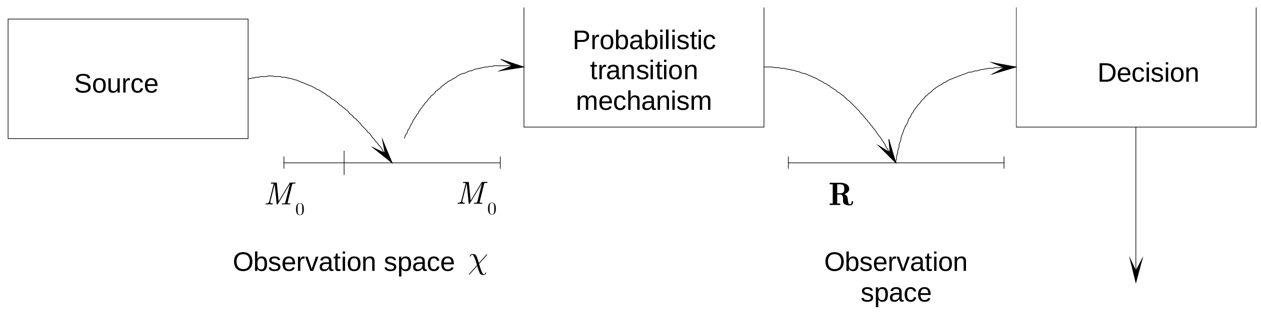
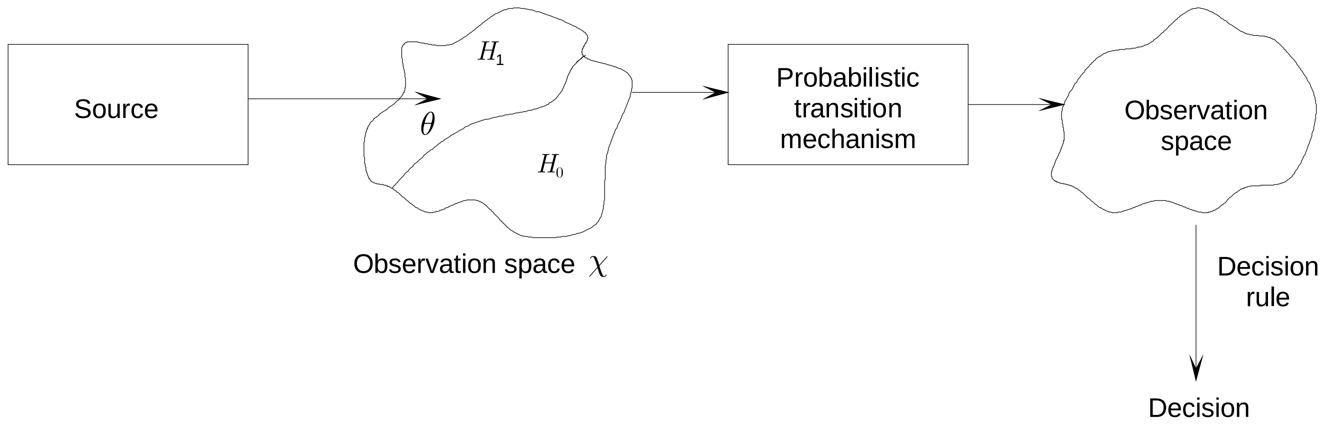
Composite Hypotheses testing

- In many hypothesis testing problems there are many possible distributions that can occur under each of the hypotheses.
- The output of the source is a set of parameters (points in a parameter space χ).
- The hypothesis corresponds to subsets of χ.
- The probability density covering the mapping from the parameter space to the observation space is denoted by $p_{\mathrm{r}|\theta}\left(\mathrm{R} \mid \theta\right)$ and is assumed to be known for all values of θ in χ.
- The final component is a decision rule.





Composite hypothesis testing problem

Example:

- For two hypothesis the observed variable will be:

$$H_0 : p_{r|H_0}\left(R \mid H_0\right) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{R^2}{2\sigma^2}\right)$$

$$H_1 : p_{r|H_1}\left(R \mid H_1\right) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(R-M)^2}{2\sigma^2}\right), \quad M_0 \le M \le M_1 .$$

<u>Baysian formulation of the composite hypothesis testing problem</u>.

- We assume that the parameter is a random variable θ taking on the values in χ.

Random variable θ

- The known probability density on θ enables us to reduce the problem to a simple hypothesis-testing problem by integrating over θ.
- $p_{\mathrm{r}|\theta}\left(\mathbf{R} \mid \theta\right)$ is interpreted as the conditional distribution of $\mathbf{R}$ given θ.

$$\Lambda\left(\mathbf{R}\right) \triangleq \frac{p_{\mathbf{r}|H_1}\left(\mathbf{R} \mid H_1\right)}{p_{\mathbf{r}|H_0}\left(\mathbf{R} \mid H_0\right)} = \frac{\int_{\chi} p_{\mathbf{r}|\theta}\left(\mathbf{R} \mid \theta\right) p_{\theta|H_1}\left(\theta \mid H_1\right) d\theta}{\int_{\chi} p_{\mathbf{r}|\theta}\left(\mathbf{R} \mid \theta\right) p_{\theta|H_0}\left(\theta \mid H_1\right) d\theta}$$

Example

- We assume that the probability density governing m on H_1 is

$$p_{m|H_1}\left(M \mid H_1\right) = \frac{1}{\sqrt{2\pi}\sigma_m}\exp\left(-\frac{M^2}{2\sigma_m^2}\right), -\infty < M < \infty.$$

- The likelihood ratio becomes:

$$\Lambda\left(R\right) = \frac{\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\left(R-M\right)^2}{2\sigma^2}\right)\frac{1}{\sqrt{2\pi}\sigma_m}\exp\left(-\frac{M^2}{2\sigma_m^2}\right)dM}{\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{R^2}{2\sigma^2}\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta$$

- By integrating and taking logarithm

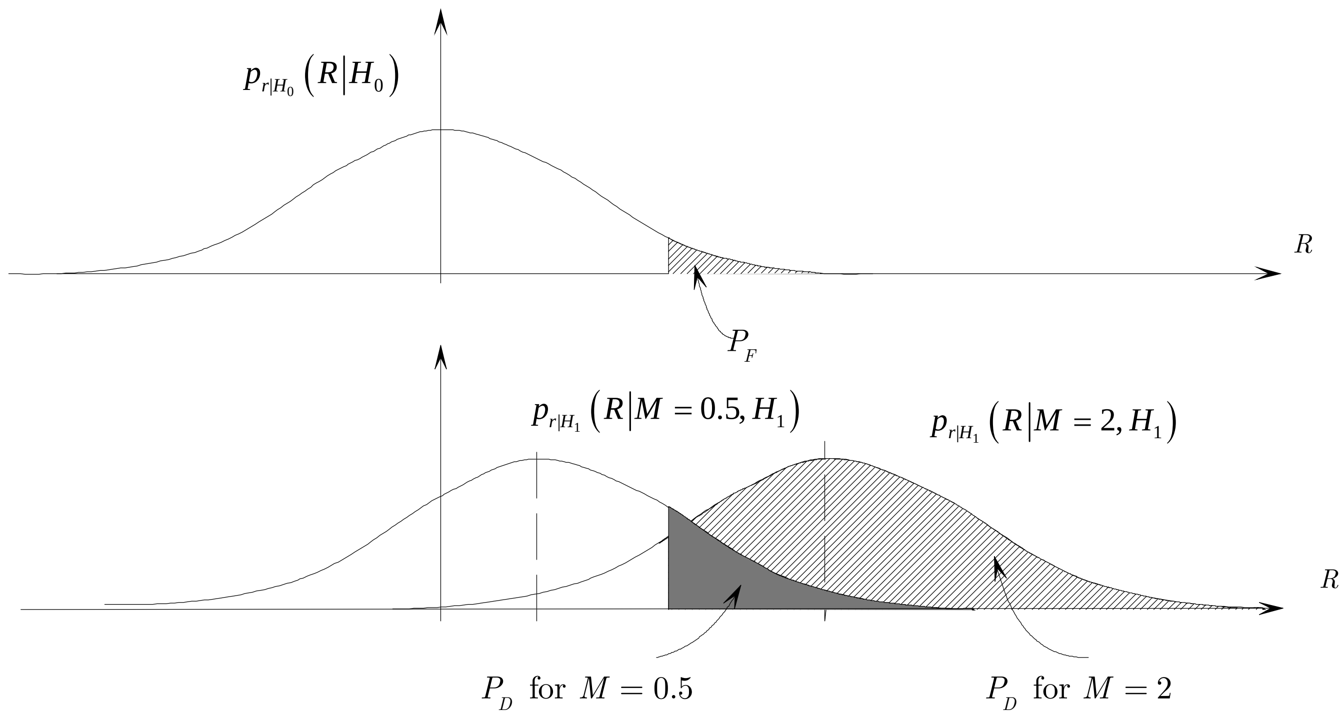
$$R^2 \underset{H_0}{\overset{H_1}{\gtrless}} \frac{2\sigma^2\left(\sigma^2+\sigma_m^2\right)}{\sigma_m^2}\left[\log\eta+\frac{1}{2}\log\left(1+\frac{\sigma_m^2}{\sigma^2}\right)\right]$$

- When θ is a random variable with an unknown density, the test procedure is not clearly specified.
  - Minimax test over the unknown density.
  - To try several densities based on partial knowledge of θ that is available. In many cases the test structure will be insensitive to the detailed behavior of the probability density.

θ nonrandom variable

- Because θ has no probability density over which to average the Bayes test in not meaningful. (We use Neyman-Pearson tests).
- Over all possible detectors that have a given P_F the one that yields the highest P_D is called Uniform Most Powerful (UMP) test.
- The best performance we could achieve would be obtained if an actual test curve equals the bound for all $M \in \chi$.
- For given P_F a uniform most powerful UMP test to exist:

  An UMP exist if we are able to design a complete likelihood ratio test (including the threshold) for every $M \in \chi$ without knowing M.
- In general the bound can be reached for any particular θ simply by designing an ordinary LRT for that particular θ.
- The UMP test must be as good as any other test for every θ.



$p_{r|H_0}\left(R|H_0\right)$
R
P_F
$p_{r|H_1}\left(R|M=0.5,H_1\right)$
$p_{r|H_1}\left(R|M=2,H_1\right)$
R
P_D for $M=0.5$
P_D for $M=2$

A necessary and sufficient condition for UMP.

A UMP test exist if and only if the likelihood ratio test for every $\theta \in \chi$ can be completely defined (including threshold) without knowledge of θ.

If UMP does not exist.

Generalized likelihood ratio test.

The perfect measurement bound suggests that a logical procedure is to estimate θ assuming H_1 is true, then estimate θ assuming H_0 is true, and use these estimates in a likelihood ratio test as if they were correct.

$$\Lambda_g(\mathrm{R}) = \frac{\max\limits_{\theta_1} p_{\mathrm{r}|\theta_1}(\mathrm{R} \mid \theta_1)}{\max\limits_{\theta_0} p_{\mathrm{r}|\theta_0}(\mathrm{R} \mid \theta_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

where θ_1 ranges over all θ in H_1 and θ_0 ranges over all θ in H_0

We make a ML estimate of θ_1, assuming that H_1 is true. We then evaluate $p_{r|\theta_1}\left(\mathrm{R} \mid \theta_1\right)$ for $\theta_1 = \hat{\theta}_1$ and use this value in numerator.

- A test contains a nuisance parameter. We are not directly concerned with the parameter it enters into the problem since it affects the PDF under H_0 and H_1.
- The GLRT decides H_1 if the fit to the data of the signal under H_1 produces a much smaller error, as measured by $\hat{\theta}_1$ than a fit to the signal under H_0 with estimated parameter $\hat{\theta}_0$

- For large data records the detector the GLRT easy to find.
- The conditions under which the asymptotic conditions hold are:

- When the data record is large and the signal is weak
  - When the Maximum Likelihood Estimation (MLE) attains it asymptotic PDF.

- The composite Hypothesis testing problem can be cast as parameter test of the PDF.
- Consider a PDF $p(\mathrm{R}, \theta)$ where θ is a $p \times 1$ vector of unknown parameters.
- The parameter test is:

$$\Lambda_g(\mathrm{R}) = \frac{p_{\mathrm{r}|\theta_1}\left(\mathrm{R}; \hat{\theta}_1, H_1\right)}{p_{\mathrm{r}|\theta_0}\left(\mathrm{R}; \theta_0, H_0\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

- Where $\hat{\theta}_1$ is the MLE of θ under H_1, the unrestricted MLE.
- $\hat{\theta}_0$ is the MLE of θ under H_0, the restricted MLE.
- As $N \to \infty$ and for unbiased estimation the variance of the estimation is given by the Cramer-Rao bound. We can express the ML estimation of the parameter $\hat{\theta}_1$ and use this value in the GLRT calculation:

Detection of Gaussially distributed random variables.

The general Gaussian problem

Hypotheses testing in case of Gaussian distribution

Equal Covariance Matrices.

Equal Mean vectors.

Definition.

- A set of random variables $r_1, r_2, \ldots, r_N$ is defined as jointly Gaussian if all their linear combinations are Gaussian random variables.
- A vector $\mathbf{r}$ is a jointly Gaussian random vector when its components are jointly Gaussian.
- In other words if

  $$z = \sum_{i=1}^{N} g_i r_i \triangleq \mathrm{G}^{\mathrm{T}} \mathrm{r}$$

  is a Gaussian random variable for all finite G^{T}, then $\mathbf{r}$ is a Gaussian vector.
- A hypothesis-testing problem is called a general Gaussian problem if $p_{\mathbf{r}|H_i}\left(\mathbf{R} \mid H_i\right)$ is a Gaussian density on all hypotheses.

- We define:

$$E(\mathbf{r}) = \mathrm{m}$$

$$\mathrm{Cov}(\mathbf{r}) = E\left[(\mathbf{r} - \mathrm{m})\left(\mathbf{r}^{\mathrm{T}} - \mathrm{m}^{\mathrm{T}}\right)\right] \triangleq \Lambda$$

$$M_{\mathrm{r}}(j\mathrm{v}) \triangleq E\left[e^{j\mathrm{v}^{\mathrm{T}}\mathbf{r}}\right] = \exp\left(j\mathrm{v}^{\mathrm{T}}\mathrm{m} - \tfrac{1}{2}\mathrm{v}^{\mathrm{T}}\Lambda\mathrm{v}\right)$$

$$p_{\mathbf{r}}(\mathbf{R}) = \left[(2\pi)^{N/2}\left|\Lambda\right|^{1/2}\right]^{-1}\exp\left[-\tfrac{1}{2}\left(\mathbf{R}^{\mathrm{T}} - \mathrm{m}^{\mathrm{T}}\right)\Lambda^{-1}(\mathbf{R} - \mathrm{m})\right]$$

- Let the observation space to be N dimensional vector (or column matrix) r:

$$\mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \end{bmatrix}$$

- Under the hypothesis H_1 we assume that $\mathbf{r}$ is a Gaussian random vector, completely specified by its mean vector and covariance matrix.

$$E\left[\mathbf{r} \mid H_1\right] = \begin{bmatrix} E\left(r_1 \mid H_1\right) \\ E\left(r_2 \mid H_1\right) \\ \vdots \\ E\left(r_N \mid H_1\right) \end{bmatrix} \triangleq \begin{bmatrix} m_{11} \\ m_{12} \\ \vdots \\ m_{1N} \end{bmatrix} \triangleq \mathrm{m}_1$$

- The covariance matrix is

$$\mathbf{K}_1 \triangleq E\left\{\left(r - m_1\right)\left(r^T - m_1^T\right)\middle|H_1\right\} = \begin{bmatrix} K_{11} & K_{12} & \cdots & K_{1N} \\ K_{21} & K_{22} & \cdots & K_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ K_{N1} & K_{N2} & \cdots & K_{NN} \end{bmatrix}$$

- The inverse of $\mathrm{K}_1 = \mathrm{Q}_1^{-1}$

  $\mathrm{K}_1\mathrm{Q}_1 = \mathrm{Q}_1\mathrm{K}_1 = \mathrm{I}$
- The probability density of $\mathbf{r}$ on H_1

  $$p_{\mathbf{r}|H_1}\left(\mathbf{R} \mid H_1\right) = \left[(2\pi)^{N/2}\left|K_1\right|^{1/2}\right]^{-1} \exp\left(-\frac{1}{2}\left(\mathbf{R}^T - \mathrm{m}_1^T\right)\mathrm{Q}_1\left(\mathbf{R} - \mathrm{m}_1\right)\right)$$
- The probability density of $\mathbf{r}$ on H_0

$$p_{\mathbf{r}|H_1}\left(\mathbf{R} \mid H_0\right) = \left[(2\pi)^{N/2}\left|K_0\right|^{1/2}\right]^{-1} \exp\left(-\frac{1}{2}\left(\mathbf{R}^T - \mathrm{m}_0^T\right)\mathrm{Q}_0\left(\mathbf{R} - \mathrm{m}_0\right)\right)$$

- Likelihood ratio test

$$\Lambda(\mathbf{R}) \triangleq \frac{p_{\mathbf{r}|H_1}\left(\mathbf{R} \mid H_1\right)}{p_{\mathbf{r}|H_0}\left(\mathbf{R} \mid H_0\right)} = \frac{\left|K_0\right|^{1/2} \exp\left(-\frac{1}{2}\left(\mathbf{R}^T - \mathrm{m}_1^T\right)\mathrm{Q}_1\left(\mathbf{R} - \mathrm{m}_1\right)\right)}{\left|K_1\right|^{1/2} \exp\left(-\frac{1}{2}\left(\mathbf{R}^T - \mathrm{m}_0^T\right)\mathrm{Q}_0\left(\mathbf{R} - \mathrm{m}_0\right)\right)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta$$

$$\frac{1}{2}\left(\mathrm{R}^T - \mathrm{m}_0^T\right)\mathrm{Q}_0\left(\mathrm{R} - \mathrm{m}_0\right) - \frac{1}{2}\left(\mathrm{R}^T - \mathrm{m}_1^T\right)\mathrm{Q}_1\left(\mathrm{R} - \mathrm{m}_1\right) \underset{H_0}{\overset{H_1}{\gtrless}}$$

$$\ln\eta + \frac{1}{2}\ln\left|\mathrm{K}_1\right| - \frac{1}{2}\ln\left|\mathrm{K}_0\right| \triangleq \gamma$$

- The test consists of finding the difference between two quadratic forms.

Special case: Equal covariance matrices.

$$\mathrm{K}_1 = \mathrm{K}_0 \triangleq \mathrm{K}.$$

$$\mathrm{Q} = \mathrm{K}^{-1}.$$

$$\left(\mathrm{m}_1^T - \mathrm{m}_0^T\right)\mathrm{Q}_0\mathrm{R} \mathop{\gtrless}_{H_0}^{H_1} \ln\eta + \frac{1}{2}\left(\mathrm{m}_1^T\mathrm{Q}\mathrm{m}_1 - \mathrm{m}_0^T\mathrm{Q}\mathrm{m}_0\right) \triangleq \gamma.$$

$$\Delta\mathrm{m} \triangleq \mathrm{m}_1^T - \mathrm{m}_0^T.$$

$$l(\mathrm{R}) \triangleq \Delta\mathrm{m}^T\mathrm{Q}\mathrm{R} \triangleq \mathrm{R}^T\mathrm{Q}\Delta\mathrm{m} \mathop{\gtrless}_{H_0}^{H_1} \gamma.$$

- $l(\mathrm{R})$ is a scalar Gaussian random variable obtained by linear transform of jointly Gaussian random variables.

- The test can completely described by the distance between the means of the two hypotheses when the variance was normalized to be equal to one.

$$d^2 \triangleq \frac{\left[E\left(l \mid H_1\right) - E\left(l \mid H_0\right)\right]^2}{\mathrm{Var}\left(\mathrm{l}|\mathrm{H}_0\right)}$$

$$E\left(l \mid H_1\right) \triangleq \Delta \mathrm{m}^T \mathrm{Q} \mathrm{m}_1$$

$$E\left(l \mid H_0\right) \triangleq \Delta \mathrm{m}^T \mathrm{Q} \mathrm{m}_0$$

$$\mathrm{Var}\left(l \mid H_0\right) = E\left\{\left[\Delta \mathrm{m}^T \mathrm{Q}\left(\mathrm{R} - \mathrm{m}_0\right)\right]\left[\left(\mathrm{R}^T - \mathrm{m}_0^T\right)\mathrm{Q}\Delta \mathrm{m}\right]\right\}$$

$$\mathrm{Var}\left(l \mid H_0\right) = \Delta \mathrm{m}^T \mathrm{Q} \Delta \mathrm{m}$$

$$d^2 = \Delta \mathrm{m}^T \mathrm{Q} \Delta \mathrm{m}$$

- The performance for the equal covariance Gaussian case is completely determined by the quadratic form.

Examples.

Case 1: Independent Components with Equal Variance.

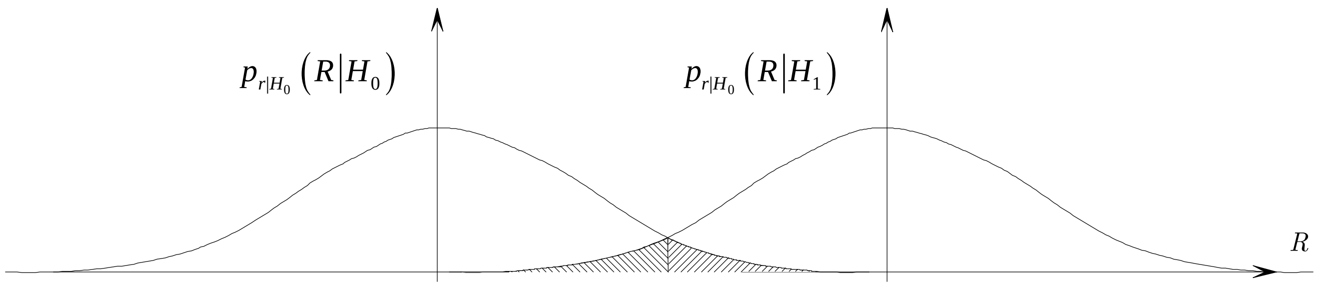
- Each r_i has the same variance σ^2 and is statistically independent:

$$\mathbf{K} = \sigma^2 \mathrm{I},$$

$$\mathbf{Q} = \frac{1}{\sigma^2} \mathrm{I},$$

- The sufficient statistics is just the dot product of the observed vector $\mathbf{R}$ and the mean difference vector $\Delta \mathrm{m}$.

$$l(\mathbf{R}) = \frac{1}{\sigma^2} \Delta \mathbf{m}^{\mathrm{T}} \cdot \mathbf{R}$$



- $d^2 = \Delta \mathrm{m}^T \frac{1}{\sigma^2} \mathrm{I} \Delta \mathrm{m} = \frac{1}{\sigma^2} \Delta \mathrm{m}^T \Delta \mathrm{m} = \frac{1}{\sigma^2} |\Delta \mathrm{m}|^2$.

  d corresponds to the distance between the two mean value vectors divided by the standard deviation of R_i.

Case 2: Independent components with Unequal Variance.

$$\mathbf{K} = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \sigma_N^2 \end{bmatrix}, \ \mathbf{Q} = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & 0 & 0 \\ 0 & \frac{1}{\sigma_2^2} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \frac{1}{\sigma_N^2} \end{bmatrix}.$$

- The sufficient statistic is $l(R) = \sum_{i=1}^{N} \frac{\Delta m_i \cdot R_i}{\sigma_i^2}$.

$$d^2 = \sum_{i=1}^{N} \frac{\left(\Delta m_i\right)^2}{\sigma_i^2}.$$

- The result can be interpreted in a new co-ordinate system

$$\Delta\mathbf{m}' = \begin{bmatrix} \frac{1}{\sigma_1} m_1 \\ \frac{1}{\sigma_2} m_2 \\ \vdots \\ \frac{1}{\sigma_N} m_N \end{bmatrix} \text{ and } \quad R_i' = \frac{1}{\sigma_i} R_i.$$

- Scale of each axis is changed so that the variances are all equal to one.
- d corresponds to the difference vector in this "scaled" coordinate system.
- In the scaled coordinate system: $\quad l(\mathbf{R}) = \Delta\mathrm{m}' \cdot \mathbf{R}_i'$.

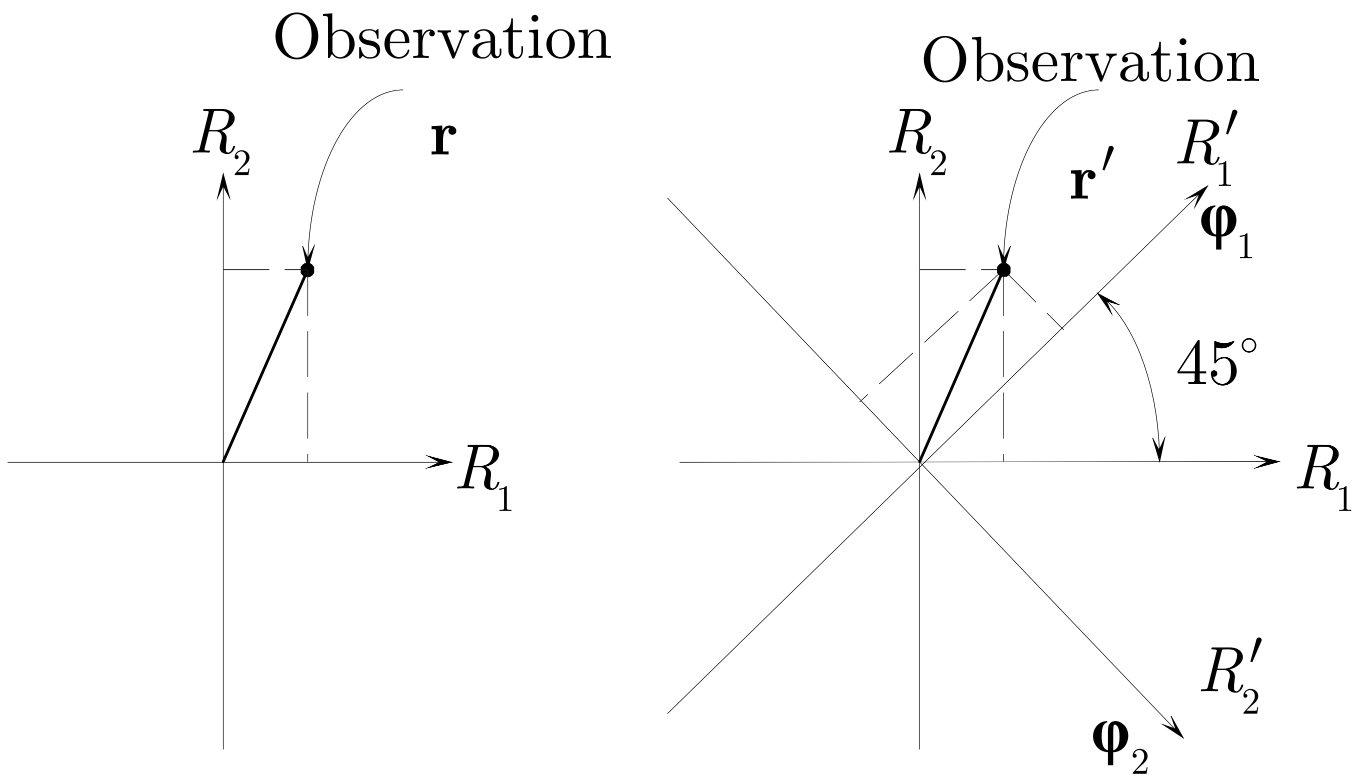
Case 3: Eigenvectors representation.

Equal mean vectors.

- We represent the R in a new coordinate system in which the components are statistically independent random variables.
- The new set of coordinate aces is defined by the orthogonal unit vectors $\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \ldots, \boldsymbol{\varphi}_N$

  $\boldsymbol{\varphi}_i^T \boldsymbol{\varphi}_j = \delta_{ij}$
- We denote the observation in the new coordinate system by $\mathbf{r}'$.
- We select the orientation of the new system so that the components r_i' and r_j' are uncorrelated.
- New component is expressed simply as a dot product:

  $r_i' = \mathbf{r}' \boldsymbol{\varphi}_i$.



Observation
r
R_2
R_1
Observation
r′
R_2
R_1'
φ_1
45°
R_1
R_2'
φ_2

- The variance matrix in the new coordinate system is calculated as $\lambda_i \delta_{ij} = \boldsymbol{\varphi}_j^T \mathbf{K} \boldsymbol{\varphi}_j$.
- The coordinate vectors should satisfy $\lambda \boldsymbol{\varphi} = \mathbf{K} \boldsymbol{\varphi}$.

Properties of the $\mathbf{K}$:

- Because $\mathbf{K}$ is symmetric, its eigenvalues are real.
- Because $\mathbf{K}$ is a covariance matrix, the eigenvalues are nonnegative.
- If the roots λ_i are distinct, the corresponding eigenvectors are orthogonal.
- If a particular root is of multiplicity M the M associated eigenvectors are linearly independent.

- The mean difference vector

$$\begin{aligned} \Delta m_1' &= \boldsymbol{\varphi}_1^T \Delta \mathbf{m} \\ \Delta m_2' &= \boldsymbol{\varphi}_2^T \Delta \mathbf{m} \\ &\vdots \\ \Delta m_N' &= \boldsymbol{\varphi}_N^T \Delta \mathbf{m} \end{aligned} .$$

- The resulting sufficient statistic in the new coordinate system is

$$l(R) = \sum_{i=1}^{N} \frac{\Delta m_i' \cdot R_i'}{\lambda_i}.$$

- There always exist a coordinate system for which the random variables are uncorrelated and that the new system is related to the old system by a linear transformation.

Equal Mean vectors.

- The mean vectors are equal

  $\mathbf{m}_i = \mathbf{m}_0 \triangleq \mathbf{m}$

$$\frac{1}{2}\left(\mathbf{R}^T - \mathbf{m}^T\right)\left(\mathbf{Q}_0 - \mathbf{Q}_1\right)(\mathbf{R} - \mathbf{m}) \underset{H_0}{\overset{H_1}{\lessgtr}} \ln\eta + \frac{1}{2}\ln\frac{\left|\mathrm{K}_1\right|}{\left|\mathrm{K}_0\right|} = \gamma$$

- The mean value vector does not contain any information telling us which of the hypothesis is true. The likelihood test subtracts them from the received vector (we may assume $\mathbf{m} = 0$).
- The difference of inverse matrices:

  $\mathbf{Q} \triangleq \mathbf{Q}_0 - \mathbf{Q}_1$

- Likelihood ratio test $l(\mathbf{R}) \triangleq \mathbf{R}^T \Delta\mathbf{Q}\mathbf{R} \underset{H_0}{\overset{H_1}{\lessgtr}} 2\gamma$

Special cases.

Case 1: Diagonal Covariance Matrix: Equal Variances.

- In case of H_1 the r_i contains the same variable as on H_0 plus additional signal components that may be correlated.

$$H_0 : r_i = n_i$$

$$H_1 : r_i = s_i + n_i$$

$$\mathbf{K}_0 = \sigma_n^2 \mathbf{I};\ \mathbf{K}_1 = \mathbf{K}_s + \sigma_n^2 \mathbf{I}$$

$$\mathbf{Q}_0 = \frac{1}{\sigma_s^2}\mathbf{I};\ \mathbf{Q}_1 = \frac{1}{\sigma_s^2}\left(\mathbf{I} + \frac{1}{\sigma_s^2}\mathbf{K}_S\right)^{-1} = \frac{1}{\sigma_s^2}(\mathbf{I} - \mathbf{H})$$

$$H = \left(\sigma_s^2 I + K_s\right)^{-1} K_s = \mathbf{Q}_0 - \mathbf{Q}_1 = \Delta\mathbf{Q}$$

$$l(\mathbf{R}) \triangleq \frac{1}{\sigma_s^2}\mathbf{R}^T \mathbf{H} \mathbf{R} \mathop{\gtrless}_{H_0}^{H_1} 2\gamma$$

Case 2: Symmetric Hypotheses, Uncorrelated Noise.

$$H_0 : \begin{matrix} r_i = s_i + n_i \\ r_i = n_i \end{matrix}$$

$$H_1 : \begin{matrix} r_i = n_i \\ r_i = s_i + n_i \end{matrix}$$

$$\mathbf{K}_0 = \begin{bmatrix} \mathbf{K}_s + \sigma_n^2 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \sigma_n^2 \mathbf{I} \end{bmatrix}$$

$$\mathbf{K}_1 = \begin{bmatrix} \sigma_n^2 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_s + \sigma_n^2 \mathbf{I} \end{bmatrix}$$

$$\Delta \mathbf{Q} = \begin{bmatrix} \frac{1}{\sigma_n^2}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \left(\mathbf{K}_s + \sigma_n^2 \mathbf{I}\right)^{-1} \end{bmatrix} - \begin{bmatrix} \left(\mathbf{K}_s + \sigma_n^2 \mathbf{I}\right)^{-1} & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sigma_n^2}\mathbf{I} \end{bmatrix}$$

$$\Delta \mathbf{Q} = \begin{bmatrix} \mathbf{H} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{bmatrix}$$

$$l(\mathbf{R}) = \frac{1}{\sigma_n^2} \mathbf{R}_1^{\mathbf{T}} \mathbf{H} \mathbf{R}_1 - \mathbf{R}_2^{\mathbf{T}} \mathbf{H} \mathbf{R}_2 \underset{H_0}{\overset{H_1}{\gtrless}} 2\gamma$$

Conclusions.

- The sufficient statistic for the general Gaussian problem is the difference between two quadratic forms

$$l(\mathbf{R}) = \frac{1}{2}\left(\mathbf{R}^T - \mathrm{m}_0^T\right)\mathrm{Q}_0\left(\mathbf{R} - \mathrm{m}_0\right) - \frac{1}{2}\left(\mathbf{R}^T - \mathrm{m}_1^T\right)\mathrm{Q}_1\left(\mathbf{R} - \mathrm{m}_1\right).$$

- A particular simple case was the one where the covariance matrixes of the hypotheses were equal. Then LLR test is $l(\mathbf{R}) = \frac{1}{2}\boldsymbol{\Delta}\mathbf{m}^{\mathrm{T}} \cdot \mathbf{Q} \cdot \mathbf{R}$.
- And the performance is characterized by $d^2 = \Delta\mathrm{m}^{\mathrm{T}} \cdot \mathbf{Q} \cdot \Delta\mathrm{m}$.
- The results described above can be obtained similarly for the M-hypothesis case.