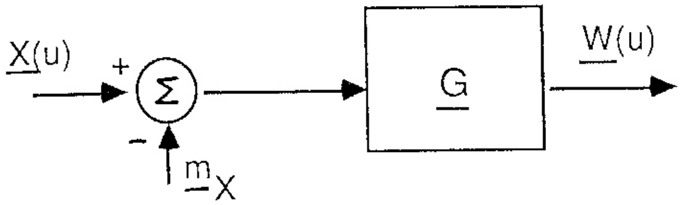


F 2-5

From the previous theory and assuming that $\underline{X}(u)$ is real, we get

$$\underline{K}_W = \underline{G}\underline{K}_X\underline{G}^T$$

and we want $\underline{K}_X = \underline{I}$. But we know that we can have the following decomposition

$$\underline{K}_X = \underline{H}\underline{H}^T$$

we want

$$\underline{G}\underline{H}\underline{H}^T\underline{G}^T = \underline{I}$$

or

$$\underline{G}\underline{H}(\underline{G}\underline{H})^T = \underline{I}$$

The simplest thing would be to chose

$$\underline{G} = \underline{H}^{-1}$$

In terms of $\underline{E}$ and $\underline{\Lambda}$ we know that

$$\underline{H} = \underline{E}\underline{\Lambda}^{1/2}\underline{U}$$

$$\begin{aligned} \Rightarrow \underline{G} = \underline{H}^{-1} &= \left(\underline{E}\underline{\Lambda}^{1/2}\underline{U}\right)^{-1} \\ &= \underline{U}^{-1}\underline{\Lambda}^{-1/2}\underline{E}^{-1} \end{aligned}$$

Now since $\underline{U}$ is orthogonal, by definition

$$\underline{U}^{-1} = \underline{U}^T$$

and also

$$\underline{E}^{-1} = \underline{E}^T$$

This holds because

$$\begin{aligned} \underline{E}^T\underline{E} &= \begin{bmatrix} \underline{e}_1{}^T \\ -- \\ \underline{e}_2{}^T \\ \vdots \\ -- \\ \underline{e}_n{}^T \end{bmatrix} [\underline{e}_1 \mid \underline{e}_2 \mid \cdots \mid \underline{e}_n] \\ &= \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} = \underline{I} \end{aligned}$$

$$\Rightarrow \boxed{\underline{G} = \underline{U}^T\underline{\Lambda}^{-1/2}\underline{E}^T}$$

2.4 The Karhunen-Loëve expansion

So far we looked at the coloring problem and the whitening problem. The solution of these two problems hinges upon one's ability to calculate $\underline{\Lambda}$ and $\underline{E}$ which are the eigenvalues and eigenvectors of $\underline{K}_X$. Let us define

$$\underline{Y}(u) \stackrel{\text{def}}{=} \underline{U}\underline{W}(u)$$

and

$$\underline{Z}(u) = \underline{\Lambda}^{1/2}\underline{Y}(u)$$

so that

$$\underline{X}(u) = \underline{E}\underline{Z}(u)$$

<u>Claim :</u>
$\underline{Y}(u)$ is also a white vector, while $\underline{Z}(u)$ has uncorrelated components but with different variances each.

<u>Proof :</u>

$$\underline{K}_Y = \underline{U}\underline{K}_W\underline{U}^T = \underline{U}\underline{U}^T = \underline{I}$$

since $\underline{W}$ is a white vector. Also

$$\begin{aligned}\underline{K}_Z &= \underline{\Lambda}^{1/2}\underline{K}_Y(\underline{\Lambda}^{1/2})^T \\ &= \underline{\Lambda}^{1/2}\underline{I}\underline{\Lambda}^{1/2} = \underline{\Lambda} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}\end{aligned}$$

$$\Rightarrow var\{Z_i(u)\} = \lambda_i \ ; \ \ i = 1, \ldots, n$$

Finally

$$\underline{X}(u) = \underline{E}\underline{Z}(u) = [\underline{e}_1 \mid \ldots \mid \underline{e}_1] \begin{bmatrix} Z_1(u) \\ \vdots \\ Z_n(u) \end{bmatrix}$$

$$\Rightarrow \underline{X}(u) = \sum_{i=1}^{n} Z_i(u)\underline{e}_i$$

$$\Rightarrow \boxed{\underline{X}(u) = \textstyle\sum_{i=1}^{n} \sqrt{\lambda_i} W_i(u)\underline{e}_i}$$

This is the so called "Karhunen-Loëve " expansion of random vectors: Each r.v. $\underline{X}(u)$ with covariance $\underline{R}_X$ can be written as a sum of orthonormal eigenvectors $\underline{e}_i$, each weighted (or multiplied) by a random variable $W_i(u)$ and further scaled by $\sqrt{\lambda_i}$. There is no guarantee that for an arbitrary orthonormal basis the projectic n of a vector will yield uncorrelated components. Here lies the usefulness of the "Karhunen-Loëve " expansion in that it provides us with that orthonormal basis upon which if we project, w› will get uncorrelated projections.
Another way of looking at it is to realize that

$$\underline{W}(u) = \begin{bmatrix} W_i(u) \\ \vdots \\ W_n(u) \end{bmatrix} = \begin{bmatrix} \lambda_1{}^{-1/2}\underline{e}_1{}^T \\ \vdots \\ \lambda_n{}^{-1/2}\underline{e}_n{}^T \end{bmatrix} \begin{bmatrix} X_i(u) \\ \vdots \\ X_n(u) \end{bmatrix}$$

$$\begin{aligned} \Rightarrow W_i(u) &= \lambda_i{}^{-1/2}\underline{e}_i{}^T\underline{X}(u) \\ &= \lambda_i{}^{-1/2}(\underline{e}_i, \underline{X}(u)) \end{aligned}$$

<u>Note :</u>

$$(\underline{e}_i, \underline{X}(u)) = \underline{e}_i{}^T\underline{X}(u) = Z_i(u)$$

where (.) denotes inner product of wo vectors. So the projection of $\underline{X}$ on the e_i's will produce components that are uncorrelated and having variance $var\,\{\underline{e}_i\underline{X}(u)\} = \lambda_i$.

$$\Rightarrow var\left\{\frac{Z_i(u)}{\sqrt{\lambda_i}}\right\} = 1$$

In summary a random vector has preferences into how it is going to be distributed in space. These preferemces are revealed by the *size* (in a stochastic sense) of its projections along the corresponding directions. The size is measured by variance. The K-L expansion basically says that any random vector can be written as the sum of n uncorrelated projections on the basis of the eigenvectors of its covariance matrix.
We next present an important application which can be found in several scientific fields such as digital communication, pattern recognition, radar, economics etc.

2.5 Binary hypothesis-testing

Suppose we observe some data forming a random vector $\underline{X}(u)$ and we would like to decide between two alternatives which we call hypotheses .

$$H_1 \;:\; \underline{X}(u) = \underline{s}_1 + \underline{N}(u)$$

$$H_0 \quad : \quad \underline{X}(u) = \underline{s}_0 + \underset{\smile}{N}(u)$$

We would like to come up with a *decision rule* which takes $\underline{X}(u)$ as input and produces a decision at the output. Under certain (academic

F2-6

) circumstances a *perfect* (errorless) detection can be made. This is called singular detection. For example:

i. N(u) = $\underline{0}$

ii. When noise is in a different *direction* than the signals:

$$\underline{s}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad , \quad \underline{s}_0 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

and assume that the second component of the noise is zero i.e.

F2-7

$$\underline{N}(u) = \begin{bmatrix} N_1(u) \\ 0 \end{bmatrix}$$

In this case $\underline{X}$ is not affected by the noise

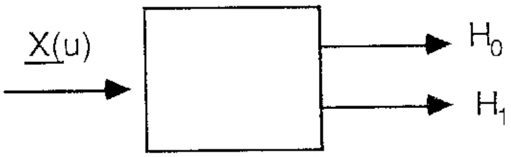
$$\underline{X}(u) = \begin{bmatrix} N_1(u) \\ 1 or - 1 \end{bmatrix}$$

iii. Bounded noise. Suppose

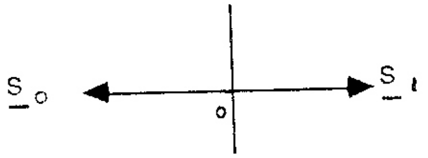
$$Pr\left[|\underline{N}(u)| \geq a\right] = 0$$

and

$$|\underline{s}_1 - \underline{s}_0| > 2a$$



F2-6



F2-7

The most familiar and useful case is when $\underline{N}(u)$ is white noise and furthermore unbounded.

<u>Note :</u>

White vectors have *no* directional preference

$$\mathcal{E}\left\{\left|\underline{N}^T(u)\underline{b}\right|^2\right\} = \underline{b}^T \underline{K}_N \underline{b} = 1$$

independent of $\underline{b}$. Assume that $\underline{K}_N = \sigma^2 \underline{I}$. A "reasonable" rule is the following: Decide hypothesis H_1 if $|\underline{X}(u) - \underline{s}_1|$ is less than $|\underline{X}(u) - \underline{s}_0|$, i.e.

$$|\underline{X}(u) - \underline{s}_1| \overset{<}{H_1} |\underline{X}(u) - \underline{s}_0|$$

In the alternative case decide H_0. This is called the minimum distance rule. It can be simplified as follows. Assume complex vectors

$$|\underline{X}(u) - \underline{s}_1| \overset{<}{H_1} |\underline{X}(u) - \underline{s}_0|$$

$$\Leftarrow |\underline{X}(u) - \underline{s}_1|^2 \overset{<}{H_1} |\underline{X}(u) - \underline{s}_0|^2$$

Expand as follows

$$\begin{aligned}
|\underline{X} - \underline{s}_1|^2 &= (\underline{X} - \underline{s}_1)^T (\underline{X} - \underline{s}_1)^* \\
&= \underline{X}^T \underline{X}^* - \underline{s}_1{}^T \underline{X}^* - \underline{X}^T \underline{s}_1{}^* + \underline{s}_1{}^T \underline{s}_1{}^* \\
&= |\underline{X}|^2 + |\underline{s}_1|^2 - 2Re\{\underline{X}^{*T} \underline{s}_1\}
\end{aligned}$$

and

$$|\underline{X} - \underline{s}_0|^2 = |\underline{X}|^2 + |\underline{s}_0|^2 - 2Re\{\underline{X}^{*T} \underline{s}_0\}$$

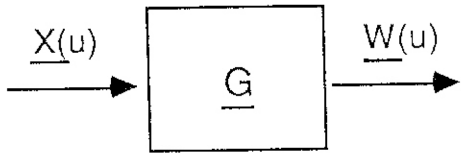
Therefore the rule reduces to

$$Re\left\{\underline{X}^{T*}(\underline{s}_1 - \underline{s}_0)\right\} \overset{>}{H_1} \frac{1}{2}\left[|\underline{s}_1|^2 - |\underline{s}_0|^2\right]$$

This decision rule corresponds to the following block diagram: . The

F2-8

basic idea is to project the vector $\underline{X}$ upon the difference of the two vectors $\underline{s}_0$ and $\underline{s}_1$ and compare it to some threshold to produce a decision.



FZ-4