

1.3 Decorelating Sets of Random Variables

- Given a random vector $\underline{u}$ of correlated components, can we transform them into a set of uncorrelated components?
Very useful, since calculations and manipulations are always easier with independent variables.

Two common ways

- eigenvector basis (Karhunen-Loève)
- Gram Schmidt

• Eigenvector basis

Consider random vector $\underline{u}$ with cov matrix $R = \overline{\underline{u}\underline{u}^T}$.

Matrix of eigenvectors of R is $Q = [\underline{q}_1 | \underline{q}_2 | \dots | \underline{q}_m]$.

Property of Hermitian R :

- real eigenvals
- eigen vecs of distinct eigenvals are orthogonal

$$\underline{q}_i^T \underline{q}_k = \delta_{ik} \quad (\text{normalized, unit length})$$

$$\text{so } Q^T Q = I \Rightarrow Q^T = Q^{-1}$$

Now express $\underline{u}$ in terms of the eigenvectors:

$$\underline{u} = \sum_{i=1}^M w_i \underline{q}_i = Q \underline{w} \quad \underline{w} = Q^{-1} \underline{u} = Q^T \underline{u}$$

random vector
of coefficients

reversible, no information lost

What is cov matrix of this new random vector? 1.3.2

$$S = \underline{w} \underline{w}^T = \underline{Q}^T R \underline{Q} = \underline{Q}^T [\lambda_1 g_1 | \dots | \lambda_M g_M] = \underline{I} \\ = \text{diag}[\lambda_1, \dots, \lambda_M]$$

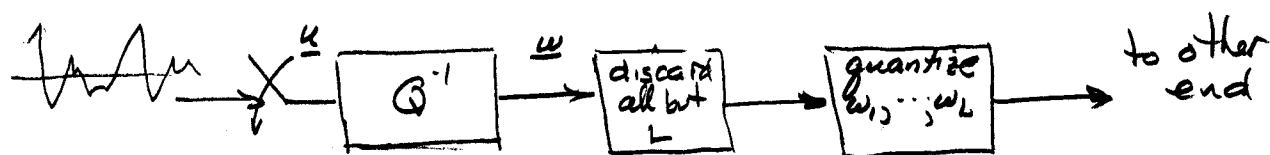
Therefore they are uncorrelated:

$$\text{cov}[w_i, w_k] = 0 \quad i \neq k$$

$$\text{var}[w_i] = \lambda_i$$

after cols of $\underline{Q}$ sorted,
so that
 $\lambda_1 > \lambda_2 > \dots > \lambda_M$

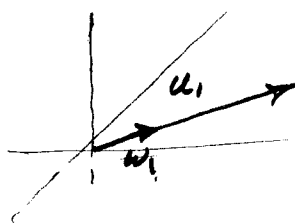
example - waveform coding



This gets rid of low variance components, ones that do not contribute much to signal. Spend bits on important components

- Gram Schmidt orthogonalization Haykin 6.7
 - This is very important in many of the filter structures used in advanced products and research.
 - It applies to any vector space (as does eigenvector basis), not just random variables.
 - We have correlated rvs $u_1, \dots, u_M$.
- Start with any one (arbitrary, but we'll use u_1).

$$\text{Normalize it } \underline{w}_1 = \frac{\underline{u}_1}{\|\underline{u}_1\|} = \frac{\underline{u}_1}{\sigma_{u_1}}$$



the first basis vector

Pick another (say, u_2)

Project it onto w_1

$$\hat{u}_2 = (u_2, w_1) w_1$$

$$= \sigma_{u_2 w_1}^2 w_1$$

Then the leftover part of u_2 is orthogonal to w_1 ,

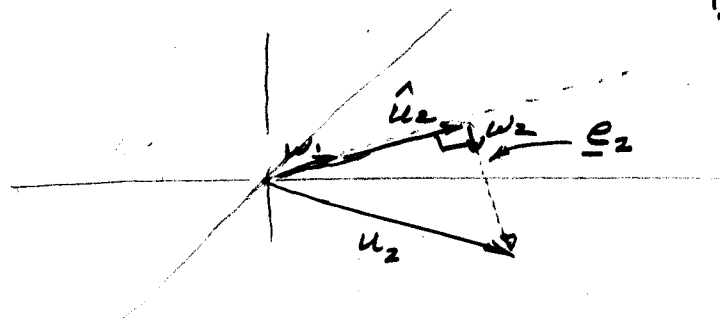
$$e_2 = u_2 - \hat{u}_2, \quad (e_2, w_1) = 0$$

$$e_2 = u_2 - \sigma_{u_2 w_1}^2 w_1$$

Normalize it

$$w_2 = \frac{e_2}{\|e_2\|} = \frac{e_2}{\sigma_{e_2}}$$

the next basis vector



Pick a third vector u_3

Project it onto the subspace spanned by w_1, w_2

$$\hat{u}_3 = (u_3, w_1) w_1 + (u_3, w_2) w_2$$

and leftover is orthog to subspace

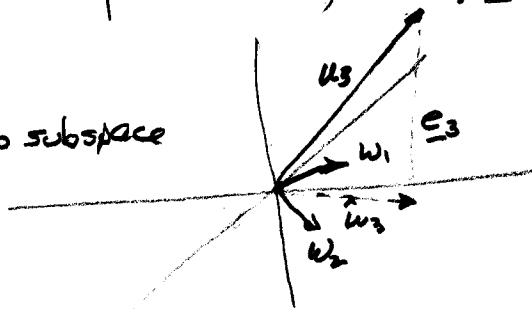
$$e_3 = u_3 - \hat{u}_3$$

$$e_3 = u_3 - \sigma_{u_3 w_1}^2 w_1 - \sigma_{u_3 w_2}^2 w_2$$

Normalize it

$$w_3 = \frac{e_3}{\|e_3\|}$$

$$w_3 = \frac{e_3}{\sigma_{e_3}}$$



and so on.

- Looking backward, write the original u_i , $i = 1..M$ in terms of the new w_i , $i = 1..M$.

1.3.4

- Define $b_{ik} = \sigma_{u_i w_k}^2$. Then

$$u_1 = b_{11} w_1$$

$$u_2 = b_{21} w_1 + b_{22} w_2$$

$\vdots$

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_M \end{bmatrix} = \begin{bmatrix} b_{11} & 0 & 0 & 0 \\ b_{21} & b_{22} & 0 & 0 \\ b_{31} & b_{32} & b_{33} & 0 \\ \vdots & \vdots & \vdots & \vdots \\ b_{M1} & b_{M2} & b_{M3} & \dots & b_{MM} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_M \end{bmatrix}$$

$$\underline{u} = B \underline{w} \quad \underline{w} = B^{-1} \underline{u} \quad B \text{ and } B^{-1} \text{ both lower triangular}$$

$$\text{Also, } R = \underline{u} \underline{u}^T = B \underline{w} \underline{w}^T B^T = B B^T = \text{Cholesky (LU) factoring}$$

- So B^{-1} decorrelates $\underline{u}$. At each stage, the estimation error is $b_{ii} w_i$ or $D \underline{w}$ where $D = \text{diag}(b_{11}, b_{22}, \dots, b_{MM})$, and the estimates are $\hat{\underline{u}} = \underline{u} - D \underline{w} = (I - D B^{-1}) \underline{u}$

Example: construct a real symmetric matrix R , perform Cholesky decomposition (Gram-Schmidt) and check that the result factors R .

$$M := 4 \quad i := 0..M-1 \quad j := 0..M-1$$

$$\text{temporary matrix: } H_{i,j} := \text{rnd}(2) - 1 \quad \text{pos. def. Hermitian: } R := H^T \cdot H$$

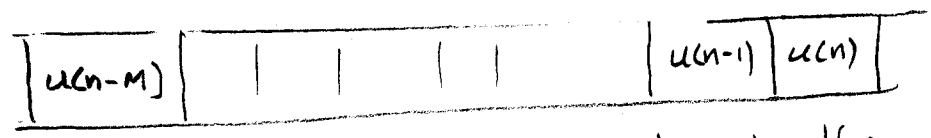
$$R = \begin{bmatrix} 1.028 & 0.514 & -0.217 & -0.717 \\ 0.514 & 0.661 & -0.034 & -0.415 \\ -0.217 & -0.034 & 0.312 & -0.356 \\ -0.717 & -0.415 & -0.356 & 1.619 \end{bmatrix} \quad \text{eigenvals}(R) = \begin{bmatrix} 0.341 \\ 0.019 \\ 0.911 \\ 2.348 \end{bmatrix}$$

$$L := \text{cholesky}(R) \quad L = \begin{bmatrix} 1.014 & 0 & 0 & 0 \\ 0.507 & 0.636 & 0 & 0 \\ -0.214 & 0.116 & 0.503 & 0 \\ -0.707 & -0.088 & -0.989 & 0.364 \end{bmatrix}$$

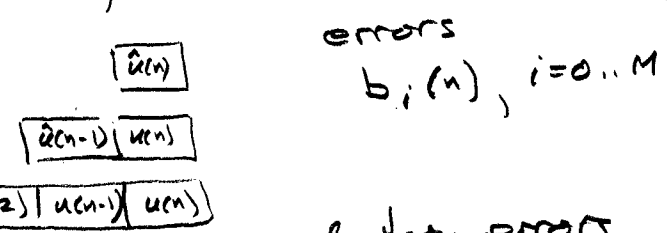
check the factorization:

$$L \cdot L^T = \begin{bmatrix} 1.028 & 0.514 & -0.217 & -0.717 \\ 0.514 & 0.661 & -0.034 & -0.415 \\ -0.217 & -0.034 & 0.312 & -0.356 \\ -0.717 & -0.415 & -0.356 & 1.619 \end{bmatrix} \quad L^{-1} = \begin{bmatrix} 0.986 & 0 & 0 & 0 \\ -0.786 & 1.574 & 0 & 0 \\ 0.601 & -0.364 & 1.989 & 0 \\ 3.363 & -0.607 & 5.409 & 2.75 \end{bmatrix}$$

Example Sequence of samples at time n

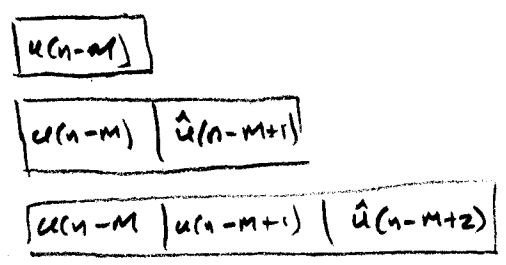


Backward prediction: start GS at $u(n)$, then work backwards



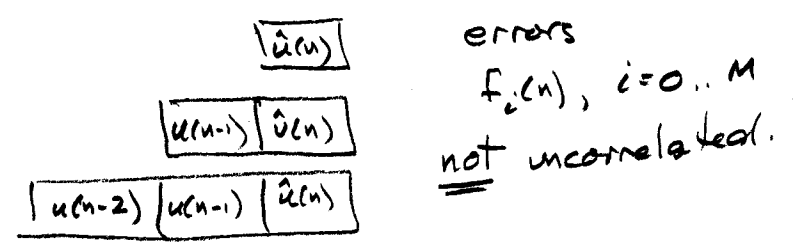
We build a set of backward predictions and prediction errors of increasing order. By GS construction, the backward pred. errors of diff orders are uncorrelated.

Forward



this also produces uncorrelated error sequence.

but by convention, the forward prediction errors of various orders at time n are all associated with prediction of $u(n)$



However

