APPENDIX Q

GRAM-SCMIDT ORTHOGONALIZATION

- From ENSC 810 Statistical Signal Processing class notes
- G-S is couched in statistical terms here, but the method applies to any vector space

1.3 Decorrelating 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 $R = \overline{\underline{u}\underline{u}^{\dagger}}$.

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

  Property of Hermitian R:
  - real eigenvals
  - eigen vecs of distinct eigenvals are orthogonal

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

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

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

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

  random vector of coefficients

  reversible, no information lost

What is cov matrix of this new random vector?

$$S = \overline{\underline{w}\,\underline{w}^T} = Q^+ R Q = Q^+[\lambda_1 q_1 | \cdots \lambda_M q_M] = \Lambda$$
$$= \mathrm{diag}[\lambda_1, \cdots \lambda_M]$$

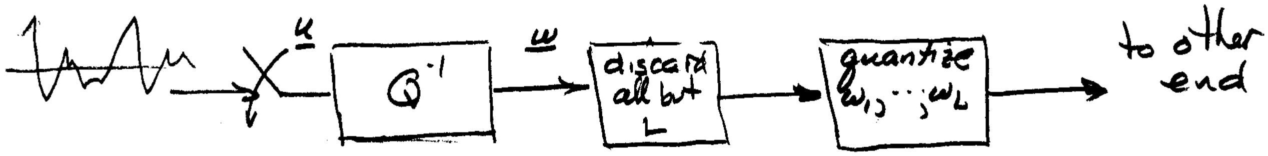
Therefore they are uncorrelated:

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

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

often cols of Q sorted, so that $\lambda_1 > \lambda_2 > \cdots > \lambda_M$

<u>example</u> – waveform coding

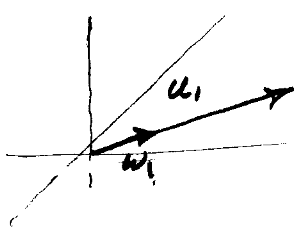


This gets rid of low variance components, ones that do not contribute much to signal. Spend bits on important components

- Gram Schmidt orthogonalization $\boxed{\text{Haykin 6.7}}$
  - 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 $u_1, \cdots u_M$.

  Start with any one (arbitrary, but we'll use u_1).

  Normalize it $\underline{w}_1 = \dfrac{u_1}{\|u_1\|} = \dfrac{u_1}{\sigma_{u1}}$



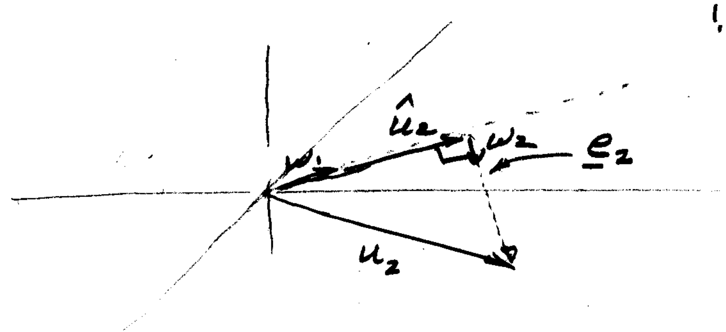
the first basis vector

Pick another (say, u_2)

Project it onto $\underline{w}_1$

$$\hat{\underline{u}}_2 = (\underline{u}_2, \underline{w}_1)\,\underline{w}_1$$

$$= \sigma^2_{u_2 w_1}\, w_1$$



Then the leftover part of $\underline{u}_2$ is orthogonal to $\underline{w}_1$,

$$\underline{e}_2 = \underline{u}_2 - \hat{\underline{u}}_2 \,, \qquad (\underline{e}, \underline{w}_1) = 0$$

$$e_2 = u_2 - \sigma^2_{u_2 w_1} w_1$$

Normalize it

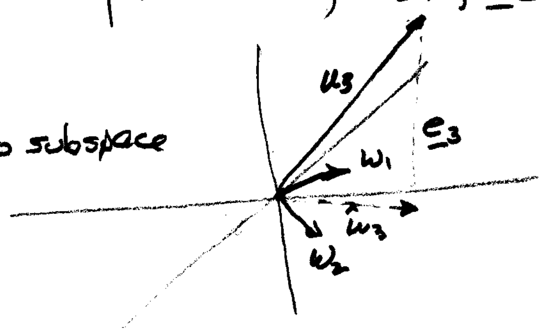
$$\underline{w}_2 = \frac{\underline{e}_2}{\|\underline{e}_2\|} = \frac{e_2}{\sigma_{e_2}} \qquad \text{the next basis vector}$$

Pick a third vector u_3

Project it onto the subspace spanned by $\underline{w}_1, \underline{w}_2$

$$\hat{u}_3 = (\underline{u}_3, \underline{w}_1)\,\underline{w}_1 + (\underline{u}_3, \underline{w}_2)\,\underline{w}_2$$

and leftover is orthog to subspace



$$\underline{e}_3 = \underline{u}_3 - \hat{\underline{u}}_3$$

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

Normalize it

$$\underline{w}_3 = \frac{\underline{e}_3}{\|\underline{e}_3\|} \qquad w_3 = \frac{e_3}{\sigma_{e_3}}$$

and so on.

- Looking backward, write the original u_i, $i = 1..M$ in terms of the new w_i, $i = 1..M$.
  - Define $b_{ik} = \sigma^2_{u_i w_k}$. Then

$$u_1 = b_{11} w_1$$
$$u_2 = b_{21} w_1 + b_{22} w_2$$
$$\vdots$$

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

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

Also, $R = \overline{\underline{u}\,\underline{u}^t} = B\,\overline{\underline{w}\,\underline{w}^t}\,B^t = BB^t$ = 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 $D = \mathrm{diag}(b_{11}, b_{22}, \cdots b_{MM})$, and the estimates are $\hat{\underline{u}} = \underline{u} - D\underline{w} = (I - DB^{-1})\,\underline{u}$

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

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

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

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

$\mathbf{L} := \mathrm{cholesky}(\mathbf{R})$

$$\mathbf{L} = \begin{bmatrix} 1.014 & 0 & 0 & 0 \\ 0.507 & 0.636 & 0 & 0 \\ -0.214 & 0.116 & 0.503 & 0 \\ -0.707 & -0.088 & -0.989 & 0.364 \end{bmatrix}$$

check the factorization:

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

Example Sequence of samples at time n

$$\boxed{u(n-M)}\,\boxed{}\,\boxed{}\,\boxed{}\,\boxed{}\,\cdots\,\boxed{u(n-1)}\,\boxed{u(n)}$$

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

$$\boxed{\hat{u}(n)}$$

$$\boxed{\hat{u}(n-1)}\,\boxed{u(n)}$$

$$\boxed{\hat{u}(n-2)}\,\boxed{u(n-1)}\,\boxed{u(n)}$$

errors
$b_i(n)$, $i = 0 .. M$

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

$$\boxed{u(n-M)}$$

$$\boxed{u(n-M)}\,\boxed{\hat{u}(n-M+1)}$$

$$\boxed{u(n-M}\,\boxed{u(n-M+1)}\,\boxed{\hat{u}(n-M+2)}$$

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 $u(n)$

$$\boxed{\hat{u}(n)}$$

$$\boxed{u(n-1)}\,\boxed{\hat{u}(n)}$$

$$\boxed{u(n-2)}\,\boxed{u(n-1)}\,\boxed{\hat{u}(n)}$$

errors
$f_i(n)$, $i = 0 .. M$
<u>not</u> uncorrelated.

However

$$\boxed{\hat{u}(n)}$$

$$\boxed{u(n)}\,\boxed{\hat{u}(n+1)}$$

$$\boxed{u(n)}\,\boxed{u(n+1)}\,\boxed{\hat{u}(n+2)}$$

$f_i(n+i)$
$i = 0 .. M$
are uncorrelated