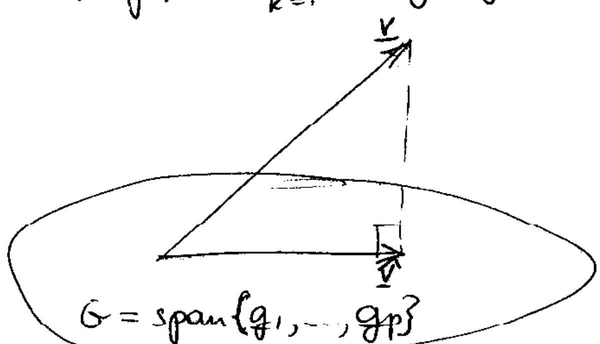
Lec 31
11/7/01 Wed

Last time: geometric viewpoint of linear prediction (and many other problems):

to find the projection
$$\hat{\underline{v}} = \sum_{k=1}^{p} \alpha_k g_k$$
of $\underline{v}$ onto span$\{g_1, \dots, g_p\}$,
need to solve the <u>normal</u> equations:
$$\langle \underline{v}, g_\ell \rangle = \sum_{k=1}^{p} \alpha_k \langle g_k, g_\ell \rangle, \quad \ell = 1, \dots, p$$



(Note: when $\{g_1, \dots, g_p\}$ is an orth. basis for G, the normal equations reduce to
$$\langle \underline{v}, g_\ell \rangle = \alpha_\ell \langle g_\ell, g_\ell \rangle$$
$$\Rightarrow \alpha_k = \frac{\langle \underline{v}, g_k \rangle}{\|g_k\|^2}, \quad k = 1, \dots, p$$
$$\Rightarrow \hat{\underline{v}} = \sum_{k=1}^{p} \frac{\langle \underline{v}, g_k \rangle}{\|g_k\|^2} g_k \, .)$$

Linear prediction:
consider zero-mean, finite-variance random variables as <u>vectors</u> is a vector space with inner product $\langle X, Y \rangle = E(XY)$.
$$\hat{S}(n) = \sum_{k=1}^{p} \alpha_k S(n-k) = \text{proj. of } S(n) \text{ onto span}\{S(n-1), S(n-2), \dots, S(n-p)\}$$
$\Rightarrow \alpha_k$'s satisfy normal eq.: $R_{SS}(\ell) = \sum_{k=1}^{p} \alpha_k R_{SS}(\ell - k)$, $k = 1, \dots, p$.

(Speech coding: partition speech into short segments; for each segment, transmit α_k's, voiced/unvoiced, and compressed error signal $\varepsilon(n) = S(n) - \hat{S}(n)$.)

Example. X, Y, W are zero-mean, statistically independent random variables, with finite and non-zero variances. Let

$$Z = X + 2Y + 3W$$

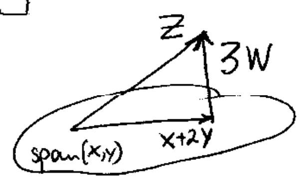
Find the minimum mean square linear estimator of Z in terms of X and Y:

find $\hat{Z} = aX + bY$, where a & b are such that $E[(Z-\hat{Z})^2] = \min$

Solution. $\hat{Z}$ = projection of Z onto $\text{span}\{X, Y\}$.

Since $W \perp \text{span}\{X, Y\}$, $\hat{Z} = X + 2Y$.

(HW 12 Prob 3 is a variation on this theme)



2.3 Recursive Estimation. [An aside]

Consider a 1-st order AR model:

$$\boxed{S(n) = a\,S(n-1) + X(n),}$$

where a is a known real number.
However, we do not observe this process $S(n)$ directly.
What we observe is $S(n)$, corrupted by additive noise
(e.g., the result of transmitting $S(n)$ over a noisy channel):

observe $Y(n) = S(n) + W(n)$;
~~we~~ would like to recover S from the observations of Y

Assume: $X(n)$ and $W(n)$ are independent of each other and indep. of $S(0)$;
$X(n)$ and $W(n)$ are ~~white~~ real-valued, zero-mean processes, with known variances λ_x and λ_w, respectively
$S(0)$ is zero-mean, with var. λ_0; indep. of $X(n)$ and $W(n)$ for $n \geq 0$

As soon as $Y(n)$ arrives, we would like to obtain:

$\hat{S}_{n|n}$ = linear least-squares estimate of $S(n)$ based on $\{Y(0), Y(1), \ldots, Y(n)\}$ [-i.e., based on all the data we have so far]
[(note that, unlike with our p-th order predictor, we use all observations, not just the last p values)];

$\lambda_{n|n} = E\{(S(n) - \hat{S}_{n|n})^2\}$ [= the corresp. mean square error – a criterion which tells us how good our estimate is;]

$\hat{S}_{n+1|n}$ = one-step linear least-squares prediction of $S(n+1)$ based on $\{Y(0), Y(1), \ldots, Y(n)\}$.

$\lambda_{n+1|n} = E\{(S(n+1) - \hat{S}_{n+1|n})^2\}$.

This looks similar to linear prediction, except that now we need to re-solve the normal equations each time a new $Y(n)$ arrives:

if $\hat{S}_{n|n} = \sum_{k=0}^{n} d_k Y(k)$, we need $S(n) - \hat{S}_{n|n} \perp Y(\ell)$ for $\ell = 0, \ldots, n \Rightarrow$
get $n+1$ normal equations: $\langle S(n), Y(\ell)\rangle = \sum_{k=0}^{n} d_k \langle Y(k), Y(\ell)\rangle$, $\ell = 0, \ldots, n$

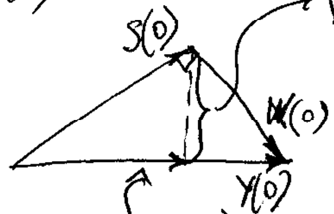
I.e., each time a new $Y(n)$ arrives, we need to solve a new $(n+1)$-by-$(n+1)$ system of linear equations! Can we compute the estimates more efficiently?

~~To get~~ Key idea: do this recursively - i.e., express the results of step n in terms of the results of step n-1.

$$\hat{S}_{0|0} = \text{proj. of } S(0) \text{ onto } Y(0)$$

$$= \frac{\langle S(0), Y(0)\rangle}{\|Y(0)\|^2} Y(0) = \frac{\langle S(0), S(0)+W(0)\rangle}{\|S(0)+W(0)\|^2} Y(0) = \frac{\|S(0)\|^2}{\|S(0)\|^2+\|W(0)\|^2} Y(0)$$

$$= \frac{\lambda_0}{\lambda_0+\lambda_w} Y(0)$$



the length of this is $\|\hat{S}_{0|0} - S(0)\| = \sqrt{\text{MS error}}$

$\hat{S}_{0|0} = \frac{\lambda_0}{\lambda_0+\lambda_w} Y(0)$

$$\lambda_{0|0} = E\left\{\left(\hat{S}_{0|0} - S(0)\right)^2\right\} = \left\| \frac{\lambda_0}{\lambda_0+\lambda_w}\left(S(0)+W(0)\right) - S(0) \right\|^2$$

$$= \left(\frac{\lambda_w}{\lambda_0+\lambda_w}\right)^2 \|S(0)\|^2 + \left(\frac{\lambda_0}{\lambda_0+\lambda_w}\right)^2 \|W(0)\|^2$$

$$\frac{\lambda_w^2\lambda_0 + \lambda_0^2\lambda_w}{(\lambda_0+\lambda_w)^2} = \frac{\lambda_0\lambda_w(\lambda_0+\lambda_w)}{(\lambda_0+\lambda_w)^2} = \frac{\lambda_0\lambda_w}{\lambda_0+\lambda_w}.$$

$$\hat{S}_{1|0} = \text{proj. of } S(1) \text{ onto } Y(0)$$

$$= \text{proj of } aS(0) \text{ onto } Y(0) + \underbrace{\text{proj. of } X(1) \text{ onto } Y(0)}_{0 \text{ since } X(1) \perp Y(0)}$$

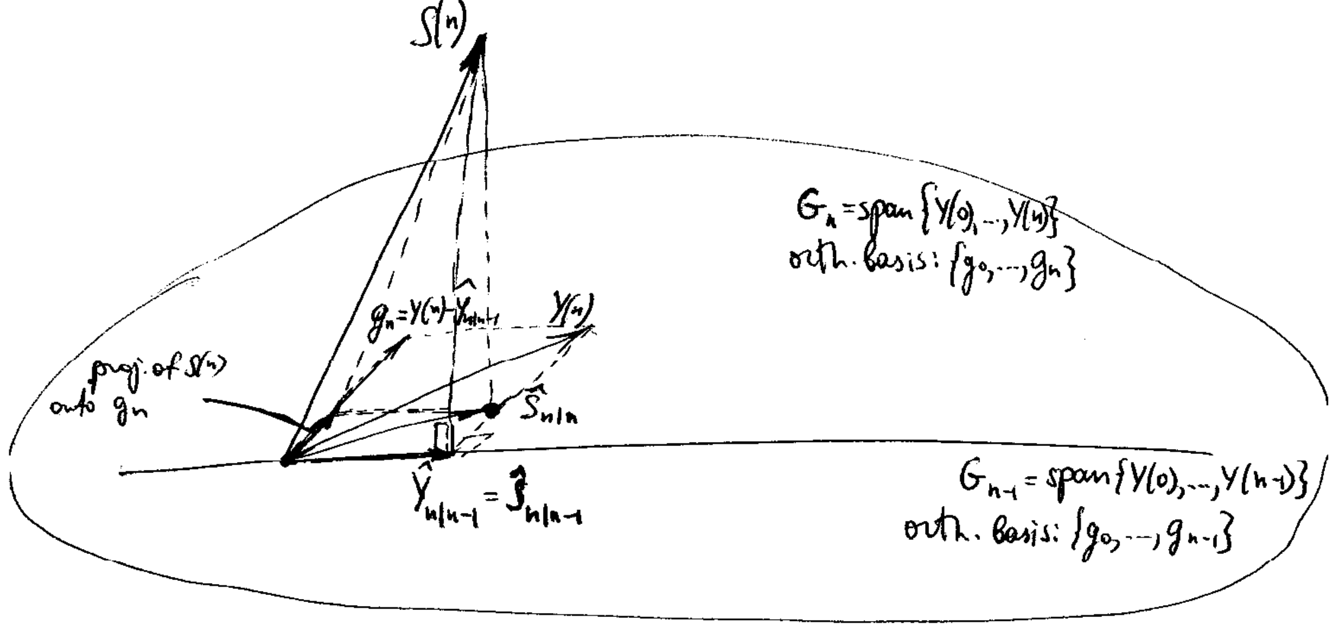
$$= a\hat{S}_{0|0}$$

$$\lambda_{1|0} = \|S_{1|0} - S(1)\|^2 = \|a\hat{S}_{0|0} - aS(0) - X(0)\|^2 =$$

$$= \|a(\hat{S}_{0|0} - S(0))\|^2 + \|X(0)\|^2$$

because $\hat{S}_{0|0} - S(0) \perp X(0)$

$$= a^2\lambda_{0|0} + \lambda_x$$



$$
\begin{aligned}
\hat{Y}_{n|n-1} &= \text{orth. proj of } Y(n) \text{ onto } G_{n-1} \\
&= \underbrace{\text{proj of } S(n) \text{ onto } G_{n-1}}_{\hat{S}_{n|n-1}} + \underbrace{\text{proj of } W(n) \text{ onto } G_{n-1}}_{\underline{0}} \\
&= \hat{S}_{n|n-1}
\end{aligned}
$$

Define $g_n = Y(n) - \hat{Y}_{n|n-1} = Y(n) - \hat{S}_{n|n-1}$

Then $g_n \perp G_{n-1}$, and $g_n \in G_n$

$\Rightarrow \{g_0,\dots,g_n\}$ is an orth. basis for G_n