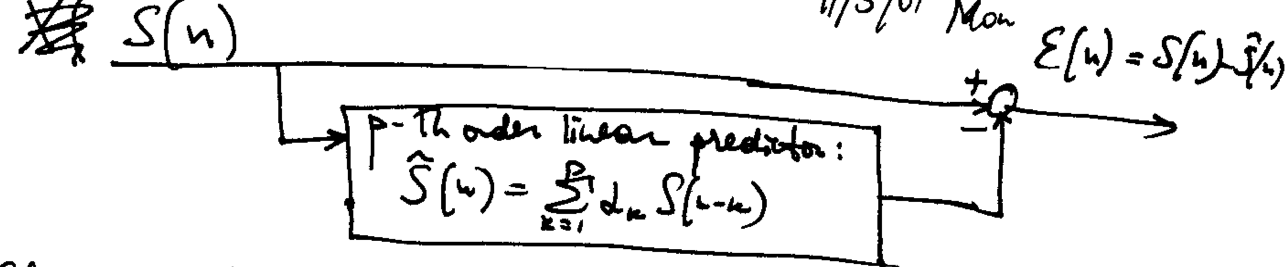
2.2 Linear Prediction

Lec 30

①

11/5/01 Mon



Choose $\alpha_1, \ldots, \alpha_p$ so as to minimize $e = E[(\varepsilon(n))^2]$

Obtained the <u>normal equations</u>:

$$r_{ss}(l) = \sum_{k=1}^{p} \alpha_k r_{ss}(l-k), \quad l = 1, \ldots, p$$

We've seen several times in this course that problems of this type often have a geometric interpretation of looking for an orthogonal projection in a vector space. This geometric interpretation is very useful, because, instead of manipulating rather complicated formulas, we can think in terms of a few rather simple and general geometric notions, such as distance, length, angle, orthogonality. In addition, we can help our intuition by drawing simple pictures in ~~2~~ a plane.

So, in order to put this problem in a geometric framework, we'll review and generalize several important concepts about vector spaces.

Geometric Interpretation of Linear Prediction. ②

Def. A vector space V is a set on which we define two operations, namely

vector addition $\underline{v}_1 + \underline{v}_2 \in V$, where $\underline{v}_1, \underline{v}_2 \in V$,

and scalar multiplication $a\underline{v} \in V$, where $\underline{v} \in V$ and $a \in \mathbb{R}$ (or $a \in \mathbb{C}$)

These operations must satisfy a set of axioms:

$\underline{u}+\underline{v} = \underline{v}+\underline{u}$ (commutativity)

$\underline{u}+(\underline{v}+\underline{w}) = (\underline{u}+\underline{v})+\underline{w}$ (associativity of +)

There exists $\underline{0} \in V$ s.t. $\underline{v}+\underline{0}=\underline{v}$ for any $\underline{v} \in V$

For any $\underline{v} \in V$, there exists $-\underline{v} \in V$ s.t. $\underline{v}+(-\underline{v})=\underline{0}$

$a(b\underline{v})=(ab)\underline{v}$ (assoc. of $\cdot$)

$$\left.\begin{array}{l} a(\underline{u}+\underline{v})=a\underline{u}+a\underline{v} \\ (a+b)\underline{v}=a\underline{v}+b\underline{v} \end{array}\right\} \text{distributivity}$$

$1 \cdot \underline{v} = \underline{v}$

Example 1. The set of all zero-mean scalar random variables is a vector space (Exercise)
I.e., each (scalar) random variable is considered a vector.
Earlier in the course, we encountered vector spaces of all deterministic, DT, N-point signals:
real-valued $\mathbb{R}^N$
and complex-valued $\mathbb{C}^N$.

Example 2. Prove that
1. The zero vector $\underline{0}$ is unique
2. For each $\underline{v}$, $-\underline{v}$ is unique
3. $0\underline{v} = \underline{0}$
4. $a\underline{0} = \underline{0}$
5. If $a\underline{v}=0$, then either $a=0$ or $\underline{v}=\underline{0}$ (or both)
6. $(-1)\underline{v} = -\underline{v}$

<u>Def.</u> A subset G of a vector space V is called a <u>subspace</u> of V if G is a vector space in its own right, i.e., if it's closed under vector addition and scalar multiplication.

<u>Def.</u> The subspace <u>spanned by</u> $g_1, \dots, g_p$ is defined by

$$\text{span}(g_1, \dots, g_p) = \left\{ \sum_{k=1}^{p} a_k g_k \text{, for all } a_1, \dots, a_p \in \mathbb{R} \text{ (or } \mathbb{C}) \right\}$$

(i.e., it's the set of all linear combinations of $g_1, \dots, g_p$.)

<u>Def.</u> The vectors $g_1, \dots, g_p$ are <u>linearly independent</u> if none of them can be represented as a linear combination of the others, i.e., if

$$\sum_{k=1}^{p} a_k g_k = 0 \quad \text{implies} \quad a_1 = a_2 = \dots = a_k$$

<u>Def.</u> The <u>dimension</u> of a subspace G is the maximum number of linearly independent vectors in G. Any such set of vectors forms a <u>basis</u> for G. (A basis necessarily <u>spans</u> G.)

<u>Def.</u> An <u>inner product</u> on a vector space V is an operation on pairs of vectors,

$$\langle u, v \rangle \in \mathbb{R} \text{ (or } \mathbb{C}), \text{ for } u, v \in V,$$

satisfying the following:

V is over $\mathbb{R}$	V is over $\mathbb{C}$
$\langle u, v \rangle = \langle v, u \rangle$	$\langle u, v \rangle = \langle v, u \rangle^*$

$$\langle au + bv, w \rangle = a\langle u, w \rangle + b\langle v, w \rangle$$

$$\langle u, u \rangle \geq 0 \text{ with } \langle u, u \rangle = 0 \text{ iff } u = 0$$

Then $\sqrt{\langle u, u \rangle}$ is called the <u>norm</u> of u, and is denoted $\|u\|$.

Example. On the space of all real-valued, zero-mean random variables, ~~$E[XY]$~~ the cross-correlation $E[XY]$ defines an inner product.

Indeed,

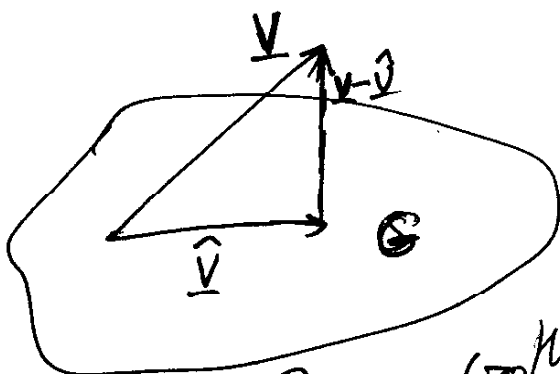
$$E[XY] = E[YX]$$

$$E[(aX+bY)Z] = aE[XZ] + bE[YZ]$$

$$E[X^2] \geq 0, \text{ with } E[X^2] = 0 \text{ iff } X = 0 \text{ with probability } 1.$$

The standard deviation $\sqrt{E[X^2]}$ is the <u>norm</u> of X.

Def. $\underline{u} \perp \underline{v}$ if $\langle \underline{u}, \underline{v} \rangle = 0$. If $\mathcal{G}$ is a subspace, $\underline{u} \perp \mathcal{G}$ means "$\underline{u}$ is orthogonal to every vector in $\mathcal{G}$. An <u>orthogonal projection</u> of $\underline{v}$ onto $\mathcal{G}$ is such vector $\hat{v}$ that $\underline{v} - \hat{\underline{v}} \perp \mathcal{G}$



Q: for random var, orthogonality means ...?

<u>Orthogonal Projection Theorem.</u> (~~Exam~~ HW 4, Problem for $\mathbb{C}^N$)
The vector $\hat{\underline{v}} \in \mathcal{G}$ minimizes $\|\underline{v} - \hat{\underline{v}}\|^2$ if and only if ~~[illegible]~~ $\hat{\underline{v}}$ is ~~an~~ the orthogonal projection of $\underline{v}$ onto $\mathcal{G}$.

Suppose ~~$S = \text{span}$~~ $\mathcal{G} = \text{span}(g_1, \dots, g_p)$. ~~form a basis for S.~~
This means that any vector in $\mathcal{G}$ can be written as a linear combination of g_k's, and, in particular,

$$\hat{\underline{v}} = \sum_{k=1}^{p} \alpha_k g_k$$

What do α_k's have to be, ~~for~~ $\hat{\underline{v}}$ to be the projections of $\underline{v}$ onto $\mathcal{G}$? Use the orthogonality $\underline{v} - \hat{\underline{v}} \perp \mathcal{G}$:

$$\underline{v} - \hat{\underline{v}} \perp g_1 \Rightarrow \langle \underline{v} - \hat{\underline{v}}, g_1 \rangle = 0$$
$$\underline{v} - \hat{\underline{v}} \perp g_2 \Rightarrow \langle \underline{v} - \hat{\underline{v}}, g_2 \rangle = 0$$
$$\vdots$$
$$\underline{v} - \hat{\underline{v}} \perp g_p \Rightarrow \langle v - \hat{\underline{v}}, g_p \rangle = 0$$

l-th equation:

$$\left\langle \underline{v} - \sum_{k=1}^{p} \alpha_k g_k, g_l \right\rangle = 0$$

$$\boxed{\langle \underline{v}, g_l \rangle = \sum_{k=1}^{p} \alpha_k \langle g_k, g_l \rangle \quad l = 1, \ldots, p}$$ ← Normal equations

<u>Example</u> Consider the space of all zero-mean real-valued random variables, with inner product

$$\langle X, Y \rangle = E[XY].$$

Let $G = \text{span}\{S(n-1), S(n-2), \ldots, S(n-p)\}$.

Find $\hat{S}(n) = \sum_{k=1}^{p} \alpha_k S(n-k)$ such that

$$\|S(n) - \hat{S}(n)\|^2 = E\left[(S(n) - \hat{S}(n))^2\right] \text{ is minimized.}$$

<u>Solution.</u> By the orthogonal projection theorem, $\hat{S}(n)$ is the projection of $S(n)$ onto G, and therefore α_k's satisfy the normal equations:

$$\langle S(n), S(n-l) \rangle = \sum_{k=1}^{p} \alpha_k \langle S(n-k), S(n-l) \rangle, \quad l = 1, \ldots, p$$

$$E[S(n) S(n-l)] = \sum_{k=1}^{p} \alpha_k E[S(n-k) S(n-l)]$$

$$r_{SS}(l) = \sum_{k=1}^{p} \alpha_k r_{SS}(l-k), \quad l = 1, \ldots, p$$

~~Solve these equations — linear algebra~~