First, Let's consider "regression" of a linear function:

$$y(t) = a + bt + n(t)$$

$$E = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} \quad x = \begin{bmatrix} a \\ b \end{bmatrix} \quad y = \begin{bmatrix} y(t_1) \\ y(t_2) \\ \vdots \\ y(t_m) \end{bmatrix}$$

$$\vec{\vec{E}}\vec{x} + \vec{n} = \vec{y} \rightarrow J = n^T n = (Ex - y)^T (Ex - y)$$

$$E^T E x = E^T y \rightarrow \boxed{\tilde{x} = (E^T E)^{-1} E^T y}$$

As $r^T q = q^T r$ if r, q are same size,

then

$$\frac{\partial (q^T r)}{\partial q} = \frac{\partial (r^T q)}{\partial q} = r$$

$$\frac{\partial (q^T q)}{\partial q} = 2q$$

And

if $J = q^T A q$

then $\frac{\partial J}{\partial q} = (A + A^T) q$

And, plugging in

$$\tilde{n} = y - E\tilde{x} = y - E(E^T E)^{-1} E^T y = (I - E(E^T E)^{-1} E^T) y$$

$$C_{\tilde{x}\tilde{x}} = \langle (\tilde{x} - x)(\tilde{x} - x)^T \rangle = (E^T E)^{-1} E^T \langle n n^T \rangle E (E^T E)^{-1}$$

$$= \sigma_n^2 (E^T E)^{-1} \quad \text{if } \langle n n^T \rangle = \sigma_2^2 \mathbb{I}$$

(uncorrelated white noise)

10 Auto Regressive Systems

AR(p)

$$S_t = \sum_{i=1}^{p} \varphi_i S_{t-i} + n_t$$

noise

parameter

Autocorrelation Fct

$$R(\tau) = \langle S\, S_{t-\tau} \rangle$$

Then $\varphi_i = R(\tau = i)$

Yule-Walker Eqtns

$$\begin{bmatrix} S_{t+1} \\ S_{t+2} \\ S_{t+3} \\ S_{t+4} \\ \vdots \\ S_{t+p} \end{bmatrix} = \begin{bmatrix} S_t & S_{t-1} & S_{t-2} & \cdots \\ S_{t+1} & S_{t-0} & S_{t-1} & \cdots \\ \vdots & \vdots & & \vdots \\ S_{t+p-1} & S_{t+p-2} & \cdots & \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_p \end{bmatrix}$$

$$S_t = \sum_{k=1}^{p} \varphi_k S_{t-k} + \sigma^2$$

Determined from above

Determined from Residuals!

Or, Frame as least squares

$$\vec{S} = \overleftrightarrow{S}_{-i} \vec{\varphi} + \vec{n}$$

$$(\vec{S} + \overleftrightarrow{S}_{-i} \vec{\varphi})^T (\vec{S} + \overleftrightarrow{S}_{-i} \vec{\varphi}) = n^T n = J$$

$$\tilde{\varphi} = (\overleftrightarrow{S}_{-i}^T \overleftrightarrow{S}_{-i})^{-1} \overleftrightarrow{S}_{-i}^T \vec{S}$$

EOFs / SVDs

Karhunen-Loeve Decomposition

not square $\rightarrow$ $E = U S V$ $\qquad S = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \\ 0 & 0 \end{bmatrix}$ or $\begin{bmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \end{bmatrix}$

In Wunsch
shown for older

$$E v_i = \lambda_i u_i$$

$$E^T u_i = \lambda_i v_i$$

$$\therefore E^T E v_i = \lambda_i^2 v_i \qquad u_i, v_i \text{ are orthonormal}$$

$$E E^T u_i = \lambda_i^2 u_i \qquad \lambda_i^2 = s_i$$

Make $U^{\in m \times m}$, $V^{\in N \times N}$ with u_i or v_i as columns, then

$$U U^T = I_m \qquad V V^T = I_N$$

$$U^T U = I_m \qquad V V^T = I_N$$

$$E V = U S \qquad E^T E V = V S^T S$$

$$E^T U = V S^T \qquad E E^T U = U S S^T$$

$$U^T E V = U^T U S \cancel{V} V^T = S$$

$$\boxed{E = U S V^T}$$

Now, We can approximate a Matrix M as

$M_{\#} = U S_{\#} V^T$ where $S_{\#}$ preserves only the $\#$ most important/largest S_i

$$M - M_{\#} = \hat{n}$$

Or, better yet

$$(M - M_{\#})^T (M - M_{\#}) = (USV^T - US_{\#}V^T)^T (USV^T - US_{\#}V^T)$$

$$= (VS^TU^T - VS_{\#}^TU^T)(USV^T - US_{\#}V^T)$$

$$= VS^TSV^T - VS_{\#}^TSV^T - VS^TS_{\#}V^T + VS_{\#}^TS_{\#}V^T$$

optimal least squares

Similarly for

$(M - M_{\#})(M - M_{\#})^T \ldots$

Consider $y = Ex + n$

$$y = USV^Tx + n$$

$y = US_{\#}V^Tx + n$, etc.