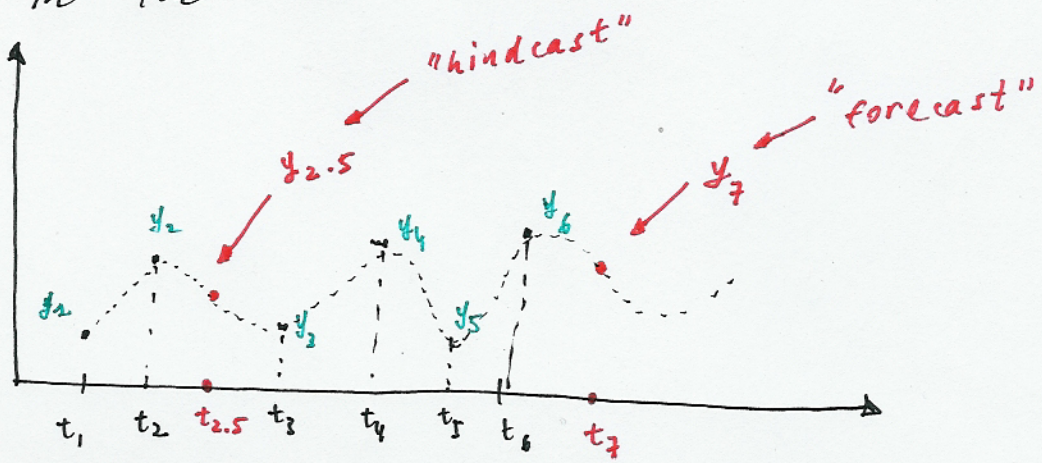
Covariance Modeling

The main idea is to use covariability between signals to interpolate, extrapolate and forecast a signal (s) of interest.

Example in the TIME DOMAIN



(Fig. 1)

Suppose you want to find the value of $y(t=t_{2.5})$. You can use the covariance $\langle y(t)\, y(t-\tau)\rangle$ to interpolate $y(t=t_{2.5})$. This operation is called "hindcast" = "time interpolation". One could use this covariance model to find the value of $y(t=t_7)$ at a future time. This operation is called "forecast" = "time extrapolation".

{interpolation / hindcast}: estimate the value of $y(t_n)$ within inside the data domain.

{extrapolation / forecast}: estimate the value of $y(t_n)$ outside the data domain.

"Univariate formulation"

$$y = \alpha x + n$$

model, error, predictor, predictand

$$\langle y \rangle = \emptyset \quad \langle x \rangle = \emptyset \quad \langle n \rangle = \emptyset$$

$$\alpha = \frac{\langle yx \rangle}{\langle xx \rangle}$$

— covariance Model also known as ~~best~~ "optimal linear estimator".

As you know this is also the least square estimate obtained by minimizing $J = \sum_{i=1}^{N} n_i^2$

Once you predict $\hat{y} = \alpha x$, one can also predict the error $\varepsilon = \sqrt{\langle n^2 \rangle} \rightarrow \hat{y} \pm \varepsilon$

where

$$\langle n^2 \rangle = \langle y^2 \rangle (1 - e^2)$$

$$e^2 = \frac{\langle xy \rangle^2}{\langle x^2 \rangle \langle y^2 \rangle}$$

correlation coefficient square = fraction of variance that is explained by the model.

The covariance modeling is a very general statement that reappears over and over in linear statistics with a wide range of applications and with different flavours.

"Multivariate Formulation"

$\underline{y}$ = vector of predictand

$\underline{x}$ = vector of predictor

$$\underline{y} = \underline{\underline{\alpha}}\,\underline{x} + \underline{n}$$

Similarly to the univariate case

$$\underline{y}\,\underline{x}^T = \underline{\underline{\alpha}}\,\underline{x}\,\underline{x}^T + \underline{n}\,\underline{x}^T$$

$$\langle \underline{y}\,\underline{x}^T \rangle = \underline{\underline{\alpha}}\,\langle \underline{x}\,\underline{x}^T \rangle$$

note that $\left(\langle \underline{y}\,\underline{x}^T \rangle\right)^T = \langle \underline{x}\,\underline{y}^T \rangle$

$$\underline{\underline{\alpha}} = \langle \underline{y}\,\underline{x}^T \rangle \langle \underline{x}\,\underline{x}^T \rangle^{-1}$$

↑ the covariance model

Once you predict $\hat{\underline{y}} = \underline{\underline{\alpha}}\,\underline{x}$ one can estimate the error $\underline{\varepsilon} = \sqrt{\text{diag}(\langle \underline{n}\,\underline{n}^T \rangle}$

$$\langle \underline{n}\,\underline{n}^T \rangle = \begin{bmatrix} \langle n_1 n_1 \rangle & \langle n_1 n_2 \rangle & \cdots \\ \langle n_2 n_1 \rangle & \ddots & \\ \vdots & & \langle n_N n_N \rangle \end{bmatrix}$$

diagonal elements are the error variance for each y_i

$$\langle \underline{n}\,\underline{n}^T \rangle = \langle (\underline{y} - \underline{\underline{\alpha}}\,\underline{x})(\underline{y} - \underline{\underline{\alpha}}\,\underline{x})^T \rangle$$

$$= \langle \underline{y}\,\underline{y}^T \rangle + \langle \underline{\underline{\alpha}}\,\underline{x}\,\underline{x}^T\,\underline{\underline{\alpha}}^T \rangle - \langle \underline{\underline{\alpha}}\,\underline{x}\,\underline{y}^T \rangle - \langle \underline{y}\,\underline{\underline{\alpha}}^T\,\underline{\underline{\alpha}}^T \rangle$$

$$\langle \underline{n}\,\underline{n}^T \rangle = \langle \underline{y}\,\underline{y}^T \rangle + \underline{\underline{\alpha}}\,\langle \underline{x}\,\underline{x}^T \rangle\,\underline{\underline{\alpha}}^T - \underline{\underline{\alpha}}\,\langle \underline{x}\,\underline{y}^T \rangle - \langle \underline{y}\,\underline{x}^T \rangle\,\underline{\underline{\alpha}}^T$$

using $\underline{\underline{\alpha}} = \langle \underline{y}\,\underline{x}^T \rangle \langle \underline{x}\,\underline{x}^T \rangle^{-1}$

$$\langle \underline{n}\,\underline{n}^T \rangle = \langle \underline{y}\,\underline{y}^T \rangle + \underbrace{\langle \underline{y}\,\underline{x}^T \rangle \langle \underline{x}\,\underline{x}^T \rangle^{-1}}_{\underline{\underline{\alpha}}} \langle \underline{x}\,\underline{x}^T \rangle \underbrace{\left[\langle \underline{y}\,\underline{x}^T \rangle \langle \underline{x}\,\underline{x}^T \rangle^{-1} \right]^T}_{\underline{\underline{\alpha}}^T}$$

$$- \underbrace{\langle \underline{y}\,\underline{x}^T \rangle \langle \underline{x}\,\underline{x}^T \rangle^{-1}}_{\underline{\underline{\alpha}}} \langle \underline{x}\,\underline{y}^T \rangle - \langle \underline{y}\,\underline{x}^T \rangle \underbrace{\left[\langle \underline{y}\,\underline{x}^T \rangle \langle \underline{x}\,\underline{x}^T \rangle^{-1} \right]^T}_{\underline{\underline{\alpha}}^T}$$

$$\langle \underline{n}\,\underline{n}^T \rangle = \langle \underline{y}\,\underline{y}^T \rangle + \langle \underline{y}\,\underline{x}^T \rangle \langle \underline{x}\,\underline{x}^T \rangle^{-1} \langle \underline{x}\,\underline{y}^T \rangle$$

cancel

$$- \langle \underline{y}\,\underline{x}^T \rangle \langle \underline{x}\,\underline{x}^T \rangle^{-1} \langle \underline{x}\,\underline{y}^T \rangle - \langle \underline{y}\,\underline{x}^T \rangle \langle \underline{x}\,\underline{x}^T \rangle^{-1} \langle \underline{x}\,\underline{y}^T \rangle$$

where $\langle \underline{y}\,\underline{x}^T \rangle^T = \langle \underline{x}\,\underline{y}^T \rangle$

$\left[\langle \underline{x}\,\underline{x}^T \rangle^{-1} \right]^T = \langle \underline{x}\,\underline{x}^T \rangle^{-1}$

$$\langle \underline{n}\,\underline{n}^T \rangle = \langle \underline{y}\,\underline{y}^T \rangle - \langle \underline{y}\,\underline{x}^T \rangle \langle \underline{x}\,\underline{x}^T \rangle^{-1} \langle \underline{x}\,\underline{y}^T \rangle$$

$$= \langle \underline{y}\,\underline{y}^T \rangle \left[1 - \underbrace{\langle \underline{y}\,\underline{y}^T \rangle^{-1} \langle \underline{y}\,\underline{x}^T \rangle \langle \underline{x}\,\underline{x}^T \rangle^{-1} \langle \underline{x}\,\underline{y}^T \rangle}_{\text{equivalent to } e^2,\ \text{this is the fraction of variance explained by model.}} \right]$$

if we define

$$\underline{\underline{C}}_{yx} = \langle \underline{y}\,\underline{x}^T \rangle$$

$$\underline{\underline{C}}_{yy} = \langle \underline{y}\,\underline{y}^T \rangle$$

$$\underline{\underline{C}}_{xx} = \langle \underline{x}\,\underline{x}^T \rangle$$

(1)

$$\hat{\underline{y}} = \underline{\underline{C}}_{yx}\,\underline{\underline{C}}_{xx}^{-1}\,\underline{x} \qquad (2)$$

$$\langle \underline{n}\,\underline{n}^T \rangle = \underline{\underline{C}}_{yy}\left[\,1 - \underline{\underline{C}}_{yy}^{-1}\,\underline{\underline{C}}_{yx}\,\underline{\underline{C}}_{xx}^{-1}\,\underline{\underline{C}}_{yx}^T\,\right]$$

$$\langle \underline{n}\,\underline{n}^T \rangle = \underline{\underline{C}}_{yy}\left[\,1 - \left(\underline{\underline{C}}_{yy}^{-1}\,\underline{\underline{C}}_{yx}\right)\left(\underline{\underline{C}}_{yx}\,\underline{\underline{C}}_{xx}^{-1}\right)^T\,\right] \qquad (3)$$

3 - steps

(1) compute covariances or prescribe a functional form of the covariance (e.g. Gaussian)

(2) Make your prediction of $\hat{\underline{y}}$ using the covariance model

$$\underline{\underline{C}}_{yx}\,\underline{\underline{C}}_{xx}^{-1} = \underline{\underline{\alpha}}$$

(note that if $\langle \underline{x}\,\underline{x}^T \rangle = \underline{\underline{I}}$ than $\underline{\underline{\alpha}} \equiv \underline{\underline{C}}_{yx}$)

(3) Estimate your error.
(note that if $\langle \underline{x}\,\underline{x}^T \rangle = \langle \underline{y}\,\underline{y}^T \rangle = 1$)

$$\langle \underline{n}\,\underline{n}^T \rangle = \left[\,1 - \underline{\underline{C}}_{yx}\,\underline{\underline{C}}_{yx}^T\right]$$

The model

$$\underline{\hat{y}} = \underline{\underline{C}}_{yx}\,\underline{\underline{C}}_{xx}^{-1}\,\underline{x}$$

Optimal Linear Estimater

$$\underline{\underline{\alpha}}$$

WHEN TO USE: if you know the stats. between x y

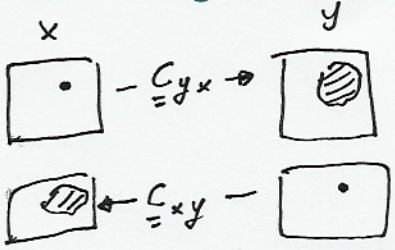
Recall the LSQ model

$$\underline{\hat{y}} = \underline{\underline{E}}\,\underline{x}$$

WHEN TO USE: if you do not know the stats. between x y you can assume a model $\underline{\underline{E}}$

this was the case where $\underline{\underline{E}}$ was choosen, $\underline{x}$ was known and $\underline{y}$ was unknown.

The analogy between $\underline{\underline{E}}$ and $\underline{\underline{C}}_{yx}\,\underline{\underline{C}}_{xx}^{-1}$ is revealing. ~~or the concept~~ When we pick a model $\underline{\underline{E}}$ to express the relationship between $\underline{y}$ and $\underline{x}$ we are really making an assumption on the covariance matrix.

"The Adjoint" $\longrightarrow$ $\underline{\underline{E}}^T = \left[\underline{\underline{C}}_{yx}\,\underline{\underline{C}}_{xx}^{-1}\right]^T = \underline{\underline{C}}_{xx}^{-1}\,\underline{\underline{C}}_{xy}$

the adjoint is an estimate of the covariance between x and y