Lecture 3

The Kalman Filter as a prime example of Outer-Inner factorization

Patrick Dewilde

TUM Inst. of Advanced Study

A quick introduction to stochastic state estimation

The basic estimation principle

A very general, universally applicable principle for signal estimation of an unknown stochastic vector *x,* given measurements contained in another stochastic vector *y* is

$$\widehat{x} = x\big|_y$$

in words: the estimate is *the (a posteriori) estimate vector "x given y "*

This property is perfectly general, whatever the distributions are, but there is a central resulting property:

Theorem: $\widehat{x}$ is such that the estimation error is uncorrelated with the observations:

$$\mathcal{E}(x-\widehat{x})y' = 0$$

in the case of zero-mean processes, this also results in minimization of the quadratic estimation error: $\widehat{x} = \arg\,\min_{w=Ky}[\mathcal{E}(x-w)(x'-w')]$ – to be worked out further (when $\mathcal{E}yy'$ is non-singular, and a linear solution $w = Ky$ is sought, this results in $K = (\mathcal{E}xy')(\mathcal{E}yy')^{-1}$.)

In the case of zero-mean Gaussian processes, this is also the maximum likelihood estimate, and it also determines the statistics of the estimate.

Estimating the new state

The classical Kalman filter starts out with knowledge of a (given) linear and time-varying system model, driven by stochastic inputs:

$$\begin{cases} x_{k+1} = A_k x_k + B_k u_k \\ y_k = C_k x_k + \nu_k \end{cases}$$

Given are: A_k, $\boldsymbol{B_k}$, $\boldsymbol{C_k}$ **(system model)**, starting at $k=0$, u_k is "system noise" on the system state x_k, and ν_k is "output noise" – We assume statistics of these signals known (actually: we shall only need second order statistics). Moreover: statistics of the input state x_0 are assumed known as well, and we assume all these signals to be independent from each other.

Therefore, let the expectations: $\mathcal{E}u_k u_k' = Q_k,\ \mathcal{E}\nu_k \nu_k' = R_k,\ \Pi_0 = \mathcal{E}x_0 x_0'$

for all $i \neq j$: $\mathcal{E}u_i u_j' = 0,\ \mathcal{E}\nu_i \nu_j' = 0,$ and for all $i.j$: $\mathcal{E}u_i \nu_j' = 0$

Asked is an (optimal) estimate $\widehat{x}_{k+1}$ of the (unknown) state x_{k+1}, given the measurements $y_0,\ y_1, \cdots,\ y_k$

Applying the principle to the Kalman filter

Preliminary remarks:

- we shall want to find recursive solutions, using the already available data at stage *k* to move to stage *k+1*.
- all processes so far are "vector" processes, correlations are matrices:

$$[\mathcal{E}xy']_{i,j} = \mathcal{E}[x_i y_j']$$

- the estimate $x|y$ is itself a stochastic vector

main property: the Kalman filter is nothing but an outer-inner factorization

the inner part takes care of the de-correlation

the outer part of the computability

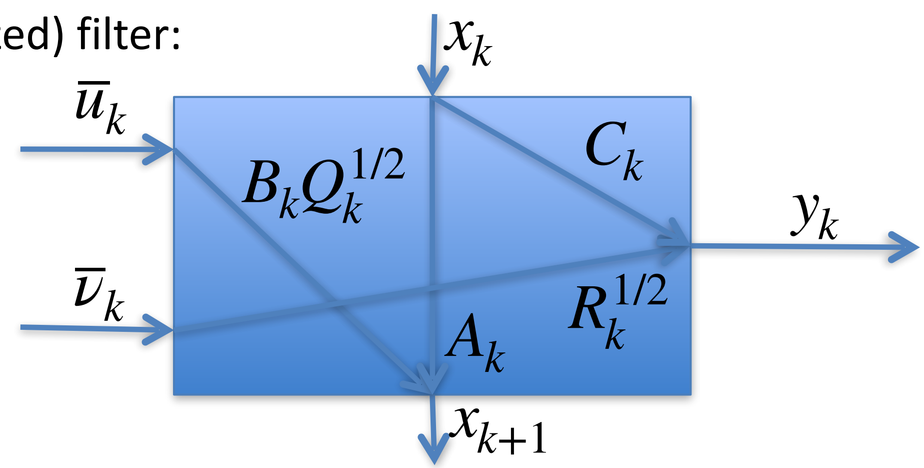
Preparatory step: normalize

$$\begin{cases} x_{k+1} = A_k x_k + \begin{bmatrix} B_k Q_k^{1/2} & 0 \end{bmatrix} \begin{bmatrix} \bar{u}_k \\ \bar{v}_k \end{bmatrix} \\ y_k = C_k x_k + \begin{bmatrix} 0 & R_k^{1/2} \end{bmatrix} \begin{bmatrix} \bar{u}_k \\ \bar{v}_k \end{bmatrix} \end{cases}$$

where now the $\overline{u}_k$ and $\overline{\nu}_k$ are fully uncorrelated with unit variance.
Let also P_k be the correlation of the innovation ($P_0=\Pi_0$ is given):

$$P_k = \mathcal{E}(x_k - \widehat{x}_k)(x_k - \widehat{x}_k)'$$

Original (normalized) filter:



Estimation as Outer-Inner filtering

First step: estimate x_1

Perform an RQ factorization (first step of the Outer-Inner algorithm):

$$\begin{bmatrix} - & - & - \\ A_0 P_0^{1/2} & B_0 Q_0^{1/2} & 0 \\ C_0 P_0^{1/2} & 0 & R_0^{1/2} \end{bmatrix} = \begin{bmatrix} - & - & - \\ 0 & M_1 & B_{o,0} \\ 0 & 0 & D_{o,0} \end{bmatrix} V_0$$

(we need to ask that P_0, Q_0, R_0 are non-singular, as well as $[A_0 \quad B_0]$ (minimal realization)

in that case $D_{o,0}$ and M_1 are square non-singular (full bases). Apply the inputs and put

$$\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix} := V_0 \begin{bmatrix} \bar{x}_0 \\ \bar{u}_0 \\ \bar{v}_0 \end{bmatrix}$$

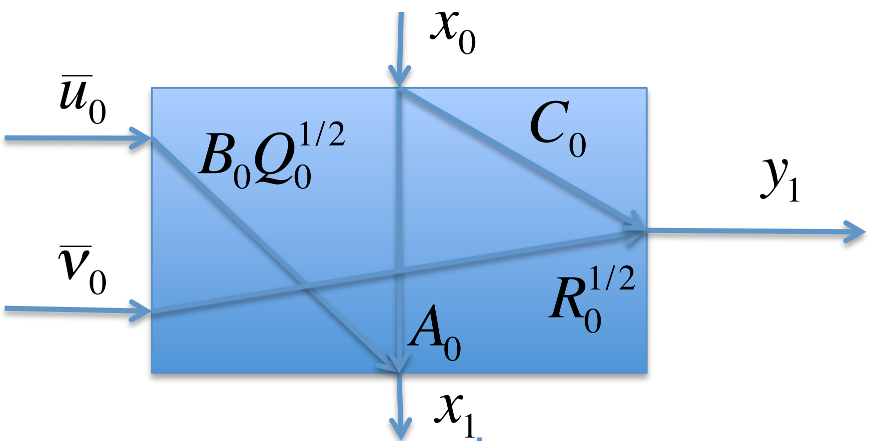
in which V_0 is unitary (it is the "Q"). Because of this the epsilons are uncorrelated and we have:

$$\begin{cases} x_1 = M_1 \epsilon_2 + B_{o,0} \epsilon_3 \\ y_0 = C_0 x_0 + v_0 = D_{o,0} \epsilon_3 \end{cases}$$

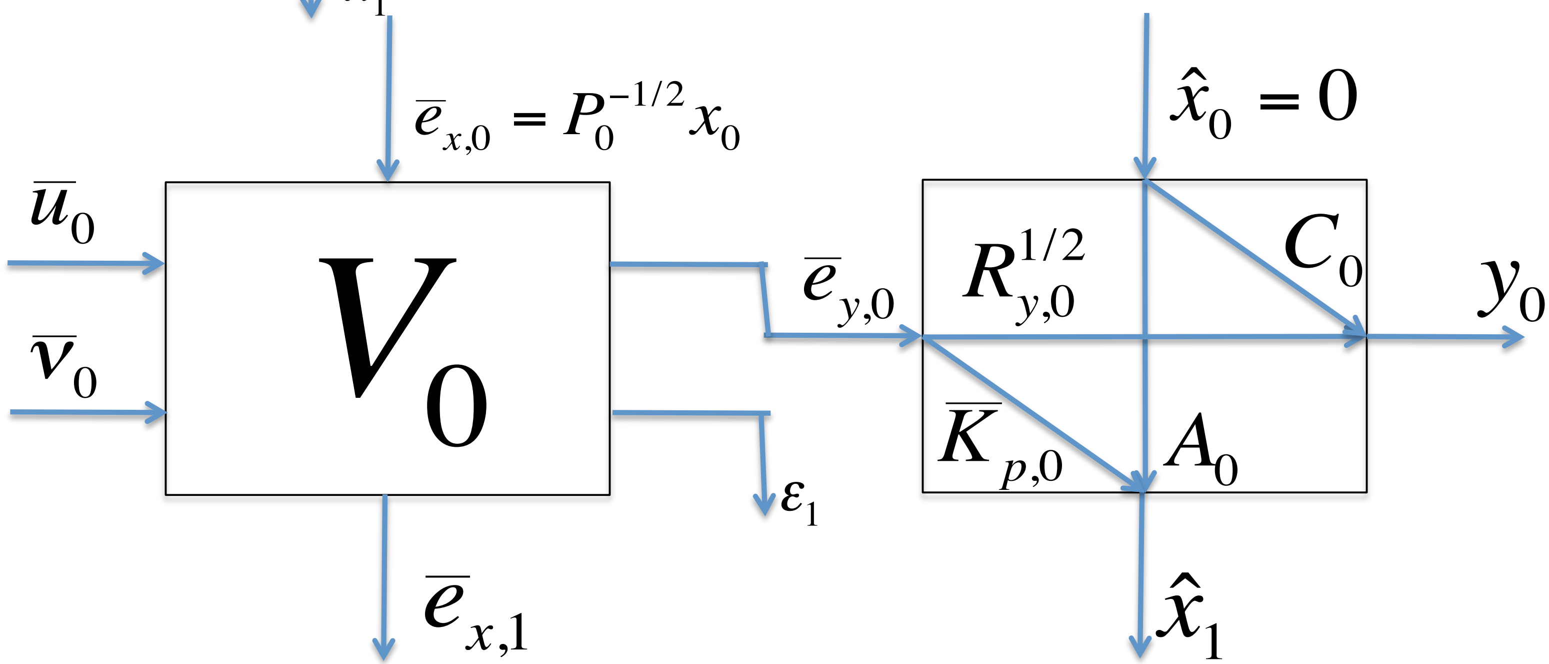
in which only ϵ_3 depends on the measurement y_0. It follows:

$$\widehat{x}_1 = B_{o.0}\epsilon_3, \;\; e_{x,1} = M\epsilon_2, P_1 = M_1$$

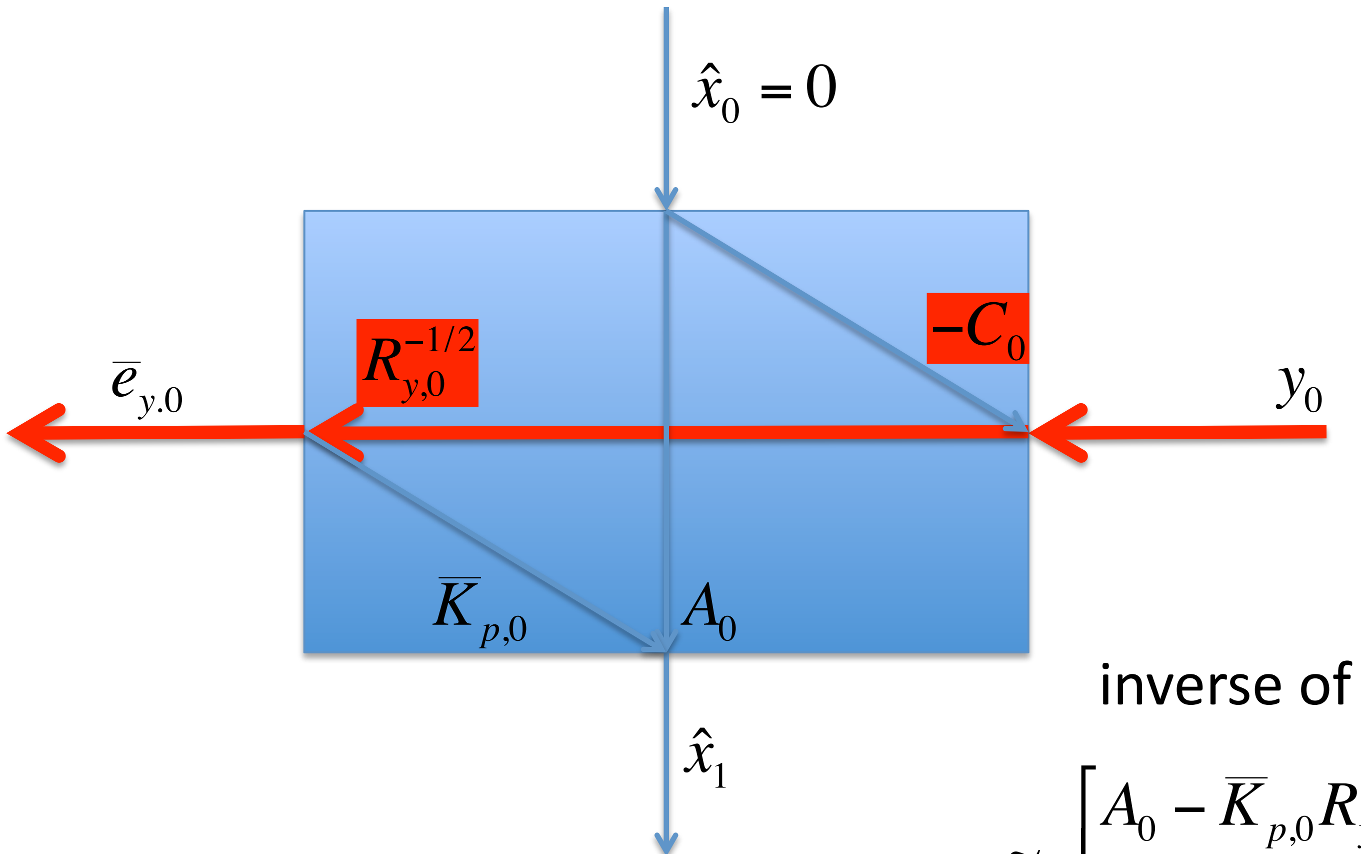
the result of the first step



becomes:



The Kalman filter for the first step



inverse of outer filter:

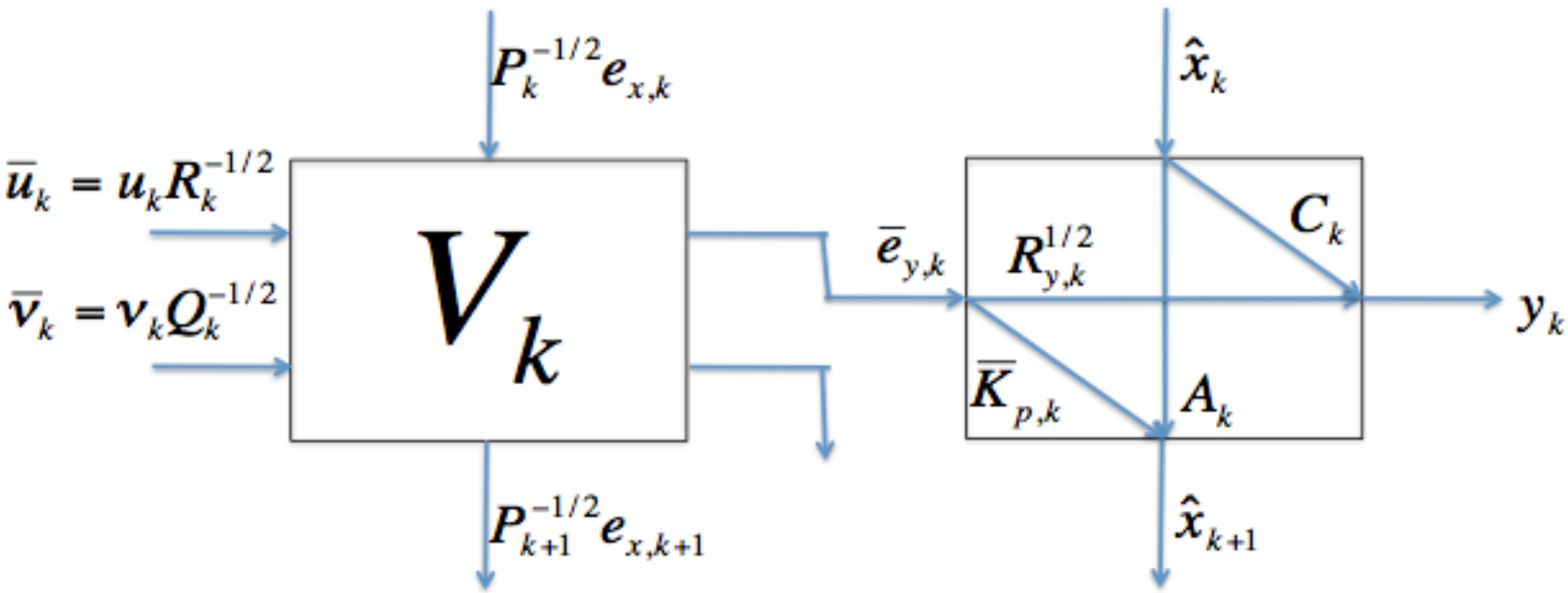
$$\sim_c \begin{bmatrix} A_0 - \bar{K}_{p,0} R_{y,0}^{-1/2} C_0 & \bar{K}_{p,0} R_{y,0}^{-1/2} \\ -R_{y,0}^{-1/2} C_0 & R_{y,0}^{-1/2} \end{bmatrix}$$

The recursive step for x_{k+1}

Again, an outer-inner factorization. Assume M_k known, calculate the kth inner and outer factors (recursive RQ):

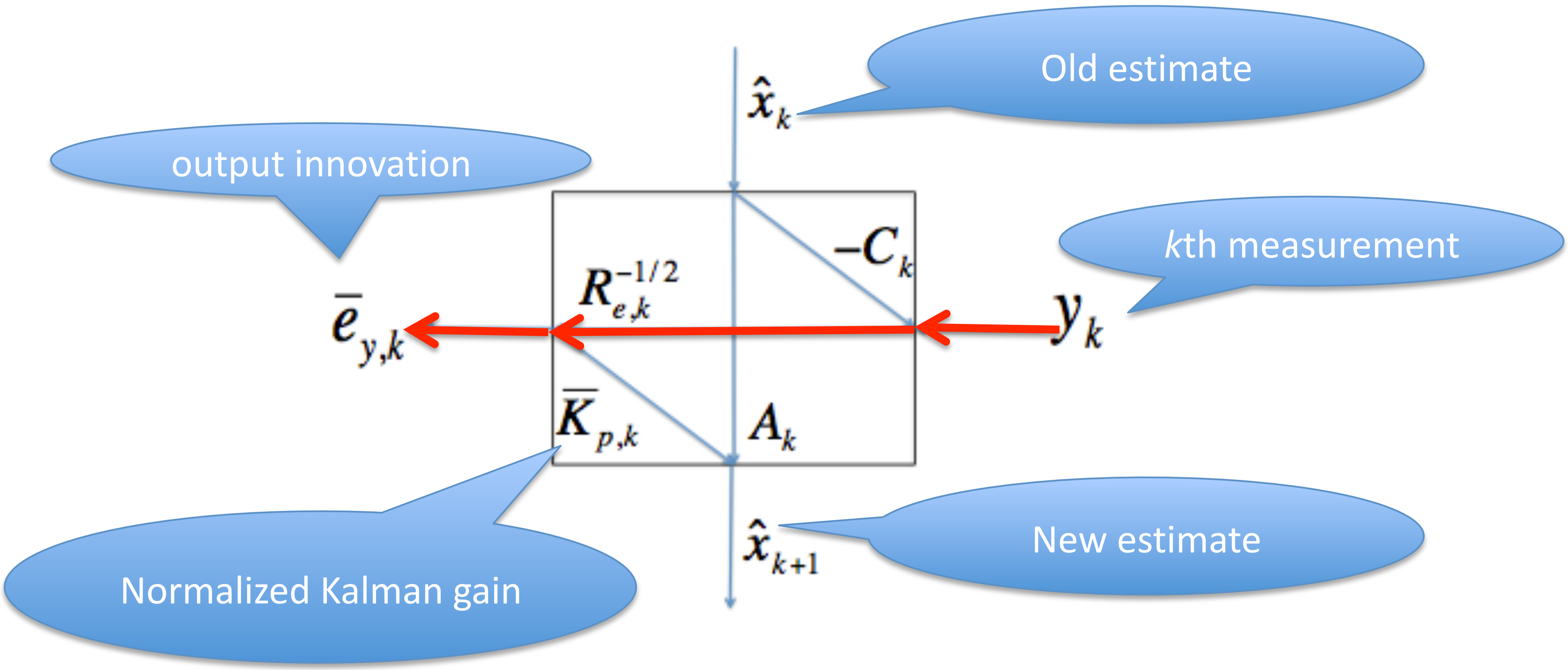
$$\begin{bmatrix} A_k M_k & B_k Q_k^{1/2} & 0 \\ C_k M_k & 0 & R_k^{1/2} \end{bmatrix} = \begin{bmatrix} 0 & M_{k+1} & B_{o,k} \\ 0 & 0 & D_{o,k} \end{bmatrix} V_k$$

Again, V_k takes care of the necessary orthogonalization, and the factor model looks as follows (soon to be proven):



The Kalman estimation filter

Is the inverse of the outer factor:

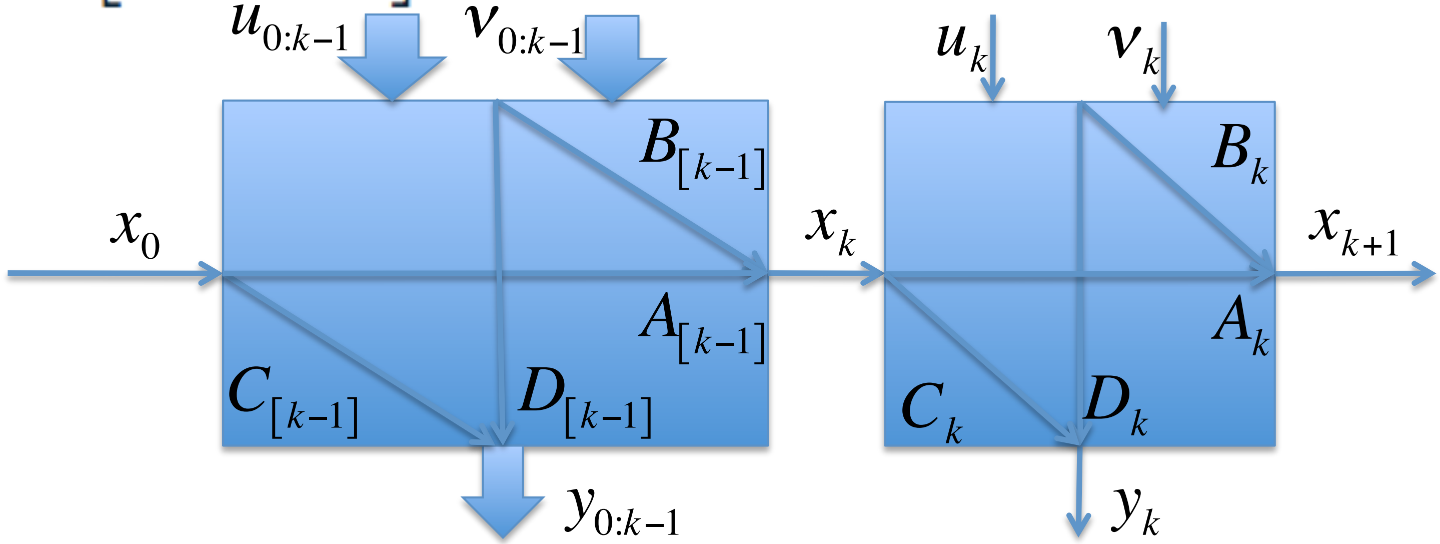


An instructive proof!

Idea: calculate the first *k* steps, as if they were the first step (bundling), then show that they unbundle as a recursion going from step *k-1* to step *k*. Let's write this bundled step with index [*k*]:

$$\begin{bmatrix} A_{[k]} & B_{[k]} \\ C_{[k]} & D_{[k]} \end{bmatrix}$$

with:



$$A_{[k]} = A_k A_{[k-1]}, B_{[k]} = \begin{bmatrix} A_k B_{[k-1]} & B_k \end{bmatrix}, C_{[k]} = \begin{bmatrix} C_{[k-1]} \\ C_k A_{[k-1]} \end{bmatrix}, D_{[k]} = \begin{bmatrix} D_{[k-1]} & \\ C_k B_{[k-1]} & D_k \end{bmatrix}$$

(and for which all original assumptions are still valid!)

Proof (cnt'd)

Hence, when an RQ at stage [*k*] looks as follows:

$$\begin{bmatrix} A_{[k]}P_0^{1/2} & B_{[k]}Q_{[k]}^{1/2} & 0 \\ C_{[k]}P_0^{1/2} & 0 & R_{[k]}^{1/2} \end{bmatrix} = \begin{bmatrix} 0 & M_{k+1} & B_{o,[k]} \\ 0 & 0 & D_{o,[k]} \end{bmatrix} \bullet V_{[k]}$$

in which, as before, $V_{[k]}$ is unitary, and we have to assume the noises to be non-singular as well as the reachability matrix $\begin{bmatrix} A_{[k]} & B_{[k]} \end{bmatrix}$, which is achieved by requesting minimality of the model. This will result in a square, non-singular M_{k+1}, and all the other results as before, in particular that $x_{k+1} - \hat{x}_{k+1} = M_{k+1}\bar{e}_{x,k+1}$.

Although this step looks global, it just summarizes the results of the local forward recursion globally.

Outer-inner recurses nicely!

We have $\begin{bmatrix} A_{[k]}M_0 & B_{[k]} \\ C_{[k]}M_0 & D_{[k]} \end{bmatrix} = \begin{bmatrix} A_k & & B_k \\ & I & \\ C_k & & D_k \end{bmatrix} \begin{bmatrix} A_{[k-1]}M_0 & B_{[k-1]} & \\ C_{[k-1]}M_0 & D_{[k-1]} & \\ & & I \end{bmatrix}$

and the Outer-Inner reduction then proceeds as follows:

first post-multiply with $\begin{bmatrix} V^*_{[k-1]} & \\ & I \end{bmatrix}$ to obtain

$$\begin{bmatrix} A_k & & B_k \\ & I & \\ C_k & & D_k \end{bmatrix} \begin{bmatrix} 0 & M_k & B_{o,[k-1]} & 0 \\ 0 & 0 & D_{o,[k-1]} & 0 \\ 0 & 0 & 0 & I \end{bmatrix} = \begin{bmatrix} 0 & A_k M_k & A_k B_{o,[k-1]} & B_k \\ 0 & 0 & D_{o,[k-1]} & 0 \\ 0 & C_k M_k & C_k B_{o,[k-1]} & D_k \end{bmatrix}$$

and then post-multiply with $\begin{bmatrix} I & & & \\ & V^*_{k,11} & & V^*_{k,21} \\ & & I & \\ & V^*_{k,12} & & V^*_{k,22} \end{bmatrix}$

Outer-inner recursion

to finally get

$$\begin{bmatrix} A_{[k]}M_0 & B_{[k]} \\ C_{[k]}M_0 & D_{[k]} \end{bmatrix} V_{[k]}^{*} = \left[\begin{array}{cc|cc} 0 & M_{k+1} & A_k B_{o,[k-1]} & B_{o,k} \\ \hline 0 & 0 & D_{o,[k-1]} & 0 \\ 0 & 0 & C_k D_{o,[k-1]} & D_{o,k} \end{array}\right]$$

as it should, with

$$V_{[k]} = \left[\begin{array}{ccc|c} I & & & \\ & V_{k,11} & & V_{k,12} \\ & & I & \\ \hline & V_{k,21} & & V_{k,22} \end{array}\right] \begin{bmatrix} V_{[k-1]} & \\ & I \end{bmatrix}$$

and

$$T_{o,[k]} = \left[\begin{array}{c|cc} A_k A_{[k-1]} & A_k B_{o,[k-1]} & B_{o,k} \\ \hline C_{[k-1]} & D_{o,[k-1]} & 0 \\ C_k A_{[k-1]} & C_k D_{o,[k-1]} & D_{o,k} \end{array}\right]$$

causally invertible!
(global Kalman filter)

Propagation of probabilities (the non-linear case)

The situation generalizes to propagation of probabilities even in the non-linear case:

non-linear model:

$$\begin{cases} x_{k+1} = F(x_k, u_k) \\ y_k = G(x_k, v_k) \end{cases}$$

We assume *u and* ν *independent processes* (not only uncorrelated), and let us now propagate probabilities (Pr), rather than expectancies.

Let

$$Y_i = \mathrm{col}[y_0, y_1, \cdots, y_i]$$

$$\widehat{x}_i = x_i|_{Y_{i-1}}, \quad f_i = x_i|_{Y_i}$$

then we show the *recursive* relations (see Kitagawa,1996 — $\mathcal{P}$ is probability):

$$\begin{cases} \widehat{x}_{k+1} & = & F(f_k, u_k) \\ \mathcal{P}(f_k) & = & \dfrac{\mathcal{P}(y_k|_{\widehat{x}_k})\mathcal{P}(\widehat{x}_k)}{\mathcal{P}(y_k)} \end{cases}$$

depends only on $\widehat{x}_k$ and y_k