$$X_n = AX_{n-1} + V_n \quad ; \quad V_n \sim \mathcal{N}(0, Q)$$

$$Y_n = CX_n + W_n \quad ; \quad W_n \sim \mathcal{N}(0, R)$$

Dimensions: $X_n, V_n$ $s \times 1$

$Y_n, W_n$ $m \times 1$

$A, Q$ $s \times s$ – "state transition matrix", "state covariance"

$C$ $m \times s$ – "measurement matrix"

$R$ $m \times m$ – "measurement covariance".

**Goal:** find $\hat{X}_n \triangleq E[X_n \mid Y_1, Y_2, \ldots Y_{n-1}]$

As with the scalar case, we use LLSE updates:

$$\hat{X}_n = A\hat{X}_{n-1} + K_n \tilde{Y}_n = \underbrace{A\hat{X}_{n-1}}_{\text{physics estimate, } s\times 1} + \underbrace{K_n}_{\text{Kalman gain, } s\times m} \underbrace{(Y_n - C\hat{X}_{n|n-1})}_{\text{innovation, } m\times 1}$$

Want $K_n$ minimizing the MSE.

Say state estimates have a covariance matrix $P_{n|n}$.

First, we make a physics prediction:

$$\hat{X}_{n|n-1} = A\hat{X}_{n-1|n-1}$$

encode the $Y_n$ in the covariance change.

$$P_{n|n-1} = AP_{n-1|n-1}A^T + Q$$

rule being used here:
$Y = AX$
$\Rightarrow \text{var } Y = A(\text{var } X)A^T$
analogous to scalar
$\text{var } y = a^2 \text{ var } x$

Then, observe $Y_n$ and update:

$$\hat{X}_{n|n} = \hat{X}_{n|n-1} + K_n(Y_n - C\hat{X}_{n|n-1})$$

$$= K_n Y_n + (I - KC)\hat{X}_{n|n-1}$$

$$P_{n|n} = K_n R K_n^T + (I-KC)P_{n|n-1}(I-KC)^T$$

To minimize MSE, set $\dfrac{\partial \operatorname{Tr} P_{n|n}}{\partial K_n} = 0$

$K_n \rightarrow m \times s$ matrix

($\operatorname{Tr} A_{n\times n} = \sum_{i=1}^{n} a_{ii}$ : minimizing $\operatorname{Tr} P_{n|n}$ is minimizing the sum of variances of each state variable)

$$\frac{\partial}{\partial K_n}\Big(\operatorname{Tr}(K_n R K_n^T) + \operatorname{Tr}(P_{n|n-1}) - 2\operatorname{Tr}(K_n C P_{n|n-1}) + \operatorname{Tr}(K_n C P_{n|n-1} C^T K_n^T)\Big) = 0$$

Trace differentiation properties: $\frac{\partial}{\partial A} \operatorname{Tr} ABA^T = A(B+B^T)$

$\frac{\partial}{\partial A} \operatorname{Tr} AC = C^T$.

$$2K_n R - 2P_{n|n-1} C^T + 2K_n C P_{n|n-1} C^T = 0$$

Solve for $K_n$:

$$\boxed{K_n = P_{n|n-1} C^T (C P_{n|n-1} C^T + R)^{-1}}$$

In steady state, $P_{n|n-1}$ and $K_n$ are constants, so

$$\boxed{K = PC^T (CPC^T + R)^{-1}}$$