# Kalman filtering is super cool!

Scalar case setup:

$$x_n = a x_{n-1} + v_n, \quad v_n \sim \mathcal{N}(0, \sigma_v^2)$$

$$y_n = b x_n + w_n, \quad w_n \sim \mathcal{N}(0, \sigma_w^2)$$

states

measurements

Want $\quad \hat{x}_n \triangleq E[x_n \mid y_1, y_2, \ldots, y_n]$

Intuition: two noisy estimates of $x_n$

Some noise is "truth", some is just from the measurement.

① The physics estimate:

given my MMSE state so far ($\hat{x}_{n-1}$)

what's going to happen next?

$$x_{n,phy} = E[x_n \mid \hat{x}_{n-1}] = a \hat{x}_{n-1}$$

② The measurement estimate:

what state corresponds most closely with the observation $y_n$?

$$x_{n,mes} = \frac{1}{b} y_n$$

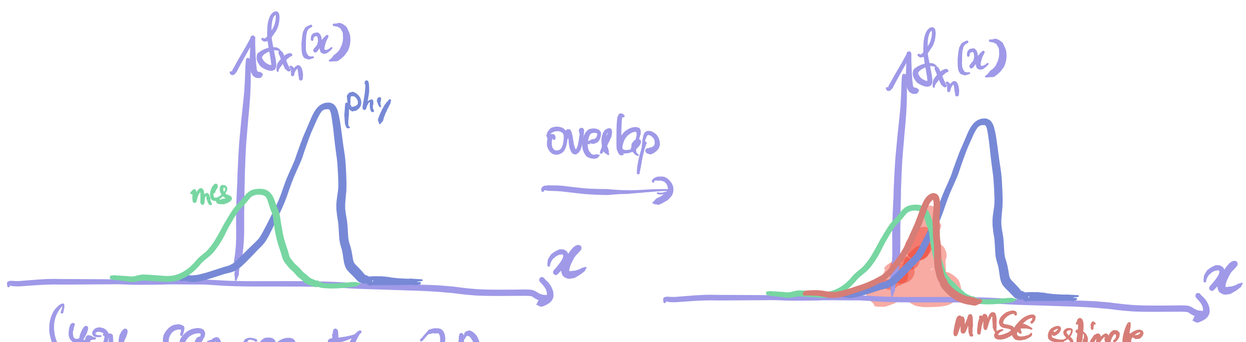
→ this doesn't exactly generalize to the vector case: with $y_n = Bx_n + w_n$, $B$ is often not square/invertible.

Want to make an overall MMSE estimate.

Intuitively, this is some linear combination of $X_{n,phy}$ and $X_{n,mes}$.

$$\hat{X}_n = \gamma X_{n,phy} + (1-\gamma) X_{n,mes}$$

Find $\gamma$ and we're done!




(you can see the 2D version of this in the KF lab!)

How do we do this overlap in algebra?

Since all variables are JG, LLSE = MMSE, so let's frame this as an LLSE update question!

(upcoming new notation: LLSE update only holds for zero-mean, so we'll say $\bar{X}_n = X_n - E[X_n]$)

Some dense algebra...

$$\hat{X}_n = E[X_n \mid Y_1, Y_2, \ldots, Y_n]$$

$$= L[X_n \mid Y_1, Y_2, \ldots, Y_n] \qquad \text{(JG)}$$

$$= L[X_n \mid Y_1] + L[\bar{X}_n \mid Y_2 - L[Y_2 \mid Y_1]] + L[\bar{X}_n \mid Y_3 - L[Y_3 \mid Y_1, Y_2]]$$
$$+ \ldots + L[\bar{X}_n \mid Y_n - L[Y_n \mid Y_1, Y_2, \ldots Y_{n-1}]]$$

$$= aL[X_{n-1} \mid Y_1] + aL[\bar{X}_{n-1} \mid Y_2 - L[Y_2 \mid Y_1]] + aL[\bar{X}_{n-1} \mid Y_3 - L[Y_3 \mid Y_1, Y_2]] \qquad \text{(LLSE update)}$$
$$+ \ldots + L[\bar{X}_n \mid Y_n - L[Y_n \mid Y_1, Y_2, \ldots, Y_{n-1}]]$$

(if you don't see $Y_n$, there's no update info, so your best guess for $X_n = a \cdot$ your best guess for $X_{n-1}$)

$$= aL[X_{n-1} \mid Y_1, Y_2, \ldots Y_{n-1}] \qquad \text{(collect all the } aL[X_{n-1} \mid \sim] \text{ terms)}$$
$$+ L[\bar{X}_n \mid Y_n - L[Y_n \mid Y_1, Y_2, \ldots, Y_{n-1}]]$$

$$= a\hat{X}_{n-1} + L[\bar{X}_n \mid Y_n - bL[X_n \mid Y_1, Y_2, \ldots, Y_{n-1}]]$$

$$= a\hat{X}_{n-1} + L[\bar{X}_n \mid Y_n - abL[X_{n-1} \mid Y_1, Y_2, \ldots, Y_{n-1}]]$$

$$= a\hat{X}_{n-1} + k_n \left(Y_n - ab\hat{X}_{n-1}\right)$$

this step from

"Kalman gain" "innovation"

$\tilde{y}_n \triangleq y_n - ab\,\hat{x}_{n-1}$

→ LLSE being linear
→ update is zero-mean

Note here:

$L[x_n | y^{(n)}] = L[x_n | y^{(n-1)}] + L[x_n | \tilde{y}_n]$

→ $k_n = \dfrac{\mathrm{cov}(x_n, \tilde{y}_n)}{\mathrm{var}\,\tilde{y}_n}$ by LLSE formula

→ innovation $\tilde{y}_n$ is zero-mean:
in expectation we have just $x_n = a^n x_0$.

How do we find $k_n$? Slightly complicated, because it's from a sequence of Gaussian overlaps:

denote a Gaussian overlap by $\otimes$

$$\hat{x}_1 \sim N(x_0, \sigma_V^2) \otimes N\left(\tfrac{1}{b} y_1, \tfrac{\sigma_W^2}{b^2}\right)$$

$$\hat{x}_2 \sim N(\hat{x}_1, \sigma_V^2) \otimes N\left(\tfrac{1}{b} y_2, \tfrac{\sigma_W^2}{b^2}\right)$$

$$= N\left(N(x_0, \sigma_V^2) \otimes N\left(\tfrac{1}{b} y_1, \tfrac{\sigma_W^2}{b^2}\right), \sigma_V^2\right) \otimes N\left(\tfrac{1}{b} y_2, \tfrac{\sigma_W^2}{b^2}\right)$$

$x_3 \sim \ldots$

This is a lot!

this was a discussion question at some point

Key here: overlap of Gaussians is Gaussian, so let's keep track of the sequence of estimator variances.

Say $\hat{X}_n \sim \mathcal{N}(X_n, \sigma_{n|n}{}^2)$.

(intuitively, e.g. $P(|\hat{x}_n - x_n| < 2\sigma_{n|n}) > 0.95$)

and say $\sigma_{n|n-1}{}^2 = a^2 \sigma_{n-1|n-1}{}^2$ – variance just of the physics estimate.

$$k_n = \frac{\operatorname{cov}(X_n, \tilde{Y}_n)}{\operatorname{var}(\tilde{Y}_n)} = \frac{\operatorname{cov}(X_n, Y_n - ab\hat{X}_{n-1})}{\operatorname{var}(\tilde{Y}_n)}$$

(I'm lazy, see website notes)

$$k_n = \frac{\sigma_{n|n-1}{}^2}{\sigma_{n|n-1}^2 + \sigma_w^2} = \frac{a^2\sigma_{n-1|n-1}{}^2}{a^2\sigma_{n-1|n-1}{}^2 + \sigma_w{}^2}$$

Lots of algebra for basically a weighted average on variances: rewrite the estimator in the form from the start.

$$\hat{X}_n = \gamma X_{n,phy} + (1-\gamma) X_{n,mes}$$

$$\hat{X}_n = a\hat{X}_{n-1} - abk_n\hat{X}_{n-1} + k_n Y_n$$

$$= (a - abk_n) X_{n,phy} + k_n b X_{n,mes}$$

This is one of the most practically applicable concepts in the course!

Example: I was in a rocketry club (stars.berkeley.edu)
We wanted to know how high the rocket went over time, because
- → knowing how high your rocket went is fun
- → it helps you time parachute release on the way down

Rocket state: $\begin{bmatrix} z \\ v_z \\ a_z \end{bmatrix}$ and sensors give you $N\left(\begin{bmatrix} z \\ a_z \end{bmatrix}, \Sigma_W\right)$.

→ Sensors are an imperfect estimate
→ We also have physics estimates that aren't perfect due to environmental noise (wind, etc) | (F=ma: the motor comes w/ a model of F(t), so we get an ideal a(t))

MMSE rocket states come from Kalman filtering w/ the physics model + sensor measurements!
(You'll work out the exact algebra for a similar case to this in the lab!)

Key to the KF equations: recursion.

$$E[Y_n \mid Y_1, \ldots, Y_{n-1}] = \alpha\beta \hat{X}_{n-1}$$

$$E[X_n \mid Y_1, \ldots, Y_{n-1}] = E[\alpha X_{n-1} + V_n \mid Y^{(n-1)}]$$
$$= \alpha \hat{X}_{n-1}$$