# Kalman filter

- Rudolf Emil Kalman — 1930 - 2016

  Stanford, then U of Florida

  gave a talk in Stuttgart while I was there but I decided not to go (mistake)

## Set-up:

$k = 0, 1, 2, \ldots$ "time"

$$x_{k+1} = M x_k \quad \leftarrow \text{Model}, \quad x_j \in \mathbb{R}^n$$

$x_0 \sim N(\mu_0, P_0)$ ← initial condition

$M \in \mathbb{R}^{n \times n}$ ← stable

This can come from a differential equation

$$\frac{dx}{dt} = f(x)$$

$$\Delta x = f(x)\,\Delta t$$

$$x_{k+1} = x_k + f(x_k)\,\Delta t$$

Suppose $f(x) = Ax$

$$x_{k+1} = (I + \Delta t A) x_k$$

$$y_{k+1} = H x_{k+1} + \eta_k \quad , \leftarrow \text{data / observations}$$

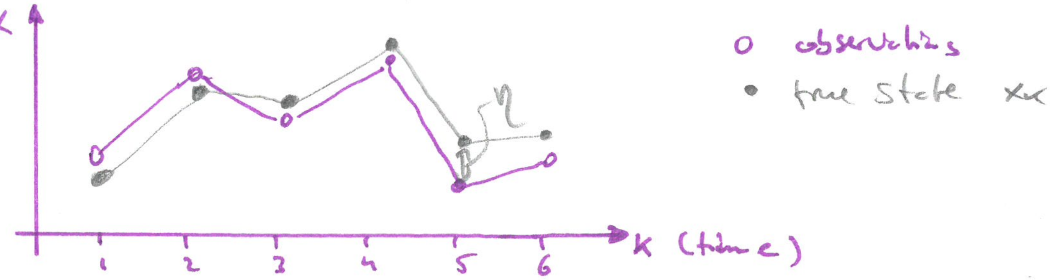
$H \in \mathbb{R}^{m \times n}$

$y \in \mathbb{R}^m$

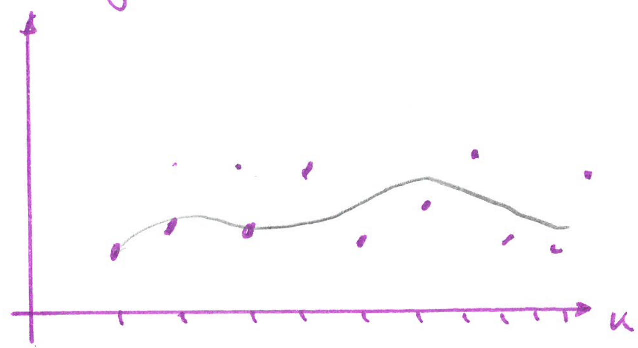
$\underbrace{\eta_k \sim N(0, R_k)}$

for now let's assume that $R_k = R$ is constant.

Cartoon:




More dramatic (larger error)




KF is a method for estimating the state of a dynamical model based on noisy observations $y$.
"Filter out the noise in $y$".

Let's consider the first step:

$$x_0 \sim \mathcal{N}(\mu_0, P_0)$$
$$x_1 = Mx_0$$
$$\underbrace{y_1 = Hx_1 + \eta_1}_{} \qquad \eta_1 \sim \mathcal{N}(0, R)$$

defines $p(y_1|x_1) = \mathcal{N}(Hx_1, R)$

$$x_1 = Mx_0 \quad , \quad x_0 \sim \mathcal{N}(\mu_0, P_0)$$
$$x_1 \sim \mathcal{N}(M\mu_0, MP_0M^T)$$

$$\underset{\text{posterior}}{p(x_1|y_1)} \propto \underset{\text{prior}}{p(x_1)}\ \underset{\text{likelihood}}{p(y_1|x_1)}$$

KF bases state estimate on this posterior distribution.

$p(x_1|y_1) \propto p(x_1)\, p(y_1|x_1)$

$$\propto \exp\left(-\tfrac{1}{2}(x_1 - M\mu_0)^T (MPM^T)^{-1}(x_1 - M\mu_0)\right)$$

$$\exp\left(-\tfrac{1}{2}(y_1 - Hx_1)^T R^{-1}(y_1 - Hx_1)\right)$$

$$\propto \exp(-J(x_1))$$

$J(x_1)$ is a quadratic!

$\Rightarrow$ $p(x_1|y_1)$ is Gaussian.

$\Rightarrow$ we only need to compute 2 parameter for this Gaussian: mean and covariance!

$\Rightarrow$ Recall HW: $x \sim p(x) = \exp(-F(x))$, $F$ is quadratic; $x \sim N(\mu, H)$
$\mu = \arg\min F$, $H$ is Hessian of $F$.

$\rightarrow$ we need to find $\arg\min J(x_1)$ and $H$.

<u>More notation:</u>

Model $\begin{cases} x_0 \sim N(\mu_0, P_0) \\ x_1 = M x_0 \quad \sim N(\underbrace{M\mu_0}_{\mu_f^b}, \underbrace{MP_0M^T}_{P_f^b}) \end{cases}$

Subscript ~~Superscript~~ $f$ is for forecast

$$J(x_1) = \frac{1}{2}(x-\mu_J)P_J^{-1}(x-\mu_J) + \frac{1}{2}(Hx-y)R^{-1}(Hx-y)$$

(dropping indices other than $J$).

$$\nabla J = P_J^{-1}(x-\mu_J) + H^T R^{-1}(Hx-y) = 0$$

$$\Rightarrow \underbrace{x^* = (P_J^{-1} + H^T R^{-1} H)^{-1}(P_J^{-1}\mu_J + H^T R^{-1} y)}_{\text{posterior mean}}$$

$$\nabla^2 J = P_J^{-1} + H^T R^{-1} H \quad \text{is Hessian}$$

$$P = (P_J^{-1} + H^T R^{-1} H)^{-1} \quad \text{is posterior covariance.}$$

$$x_1 | y_1 \sim \mathcal{N}(x^*, P)$$

Recall: Woodbury matrix identity

$$(A - BD^{-1}C)^{-1} = A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1}$$

$$(P_J^{-1} + H^T R^{-1} H)^{-1} = P_J + P_J(-H^T)(R - HP_J H^T)^{-1} H P_J$$

$A = P_J^{-1}$
$B = -H^T$
$D = R$
$C = H$

$$= P_J - P_J H^T (H P_J H^T + R)^{-1} H P_J$$

$$= (I - \underbrace{P_J H^T (HPH^T + R)^{-1}}_{K,\ \text{the Kalman gain.}} H) P_J$$

$$K = P_f H^T (H P_f H^T + R)^{-1}$$

$$\underset{\uparrow \text{ posterior covariance.}}{P} = \underbrace{(I - KH)}_{\text{update}} \underset{\uparrow \text{ forecaste covariance}}{P_f}$$

About the mean...

$$x^a = \underbrace{(P_f^{-1} + H^T R^{-1} H)^{-1}}_{(I-KH)P_f} (P_f^{-1} \mu_f + H^T R^{-1} y)$$

$$= (I - KH) P_f (P_f^{-1} \mu_f + H^T R^{-1} y)$$

$$= (I - KH) \mu_f + \underbrace{(I - KH) P_f H^T R^{-1} y}_{(P_f H^T R^{-1} - K H P_f H^T R^{-1})\, y}$$

$$= (P_f H^T R^{-1} - P_f H^T (H P_f H^T + R)^{-1} H P_f H^T R^{-1})\, y$$

$$= P_f H^T (I - (H P_f H^T + R)^{-1} H P_f H^T) R^{-1} y$$

$$= \underbrace{P_f H^T (H P_f H^T + R)^{-1}}_{K} \underbrace{(\cancel{H P_f H^T} + R - \cancel{H P_f H^T})}_{I} R^{-1} y$$

$$= K y$$

$$x^* = (I - KH)\mu_f + Ky$$

$$\boxed{x^* = \underset{\substack{\uparrow \\ \text{forecast}}}{\mu_f} + \underset{\substack{\uparrow \\ \text{gain}}}{K}\underbrace{(y - H\mu_f)}_{\text{"error"}}}$$

updated estimate.

<u>KF start to finish:</u>

$$X_0 \sim N(\mu_0, P_0)$$

forecast $\quad X_1 \sim N(\mu_f, P_f)$

observation $\quad y_1 = Hx_1 + \eta, \quad \eta \sim N(0, R)$

analysis $\quad X_a = \mu_f + K(y - H\mu_f)$

$\qquad P_a = (I - KH)P_f$

~~[illegible]~~

Next step: $p(x_2 | y_1 y_2)$

Recall the review:

$$P(A|B) = P(B|A)\frac{P(A)}{P(B)}$$
$$P(A,B|C) = P(A|B,C)\,P(B|C)$$

$$p(x_2|y_1 y_2) = p(y_1 y_2|x_2)\frac{p(x_2)}{p(y_1 y_2)}$$

$$= p(y_2|x_2)\,p(y_1|x_2)\,p(x_2)\frac{1}{p(y_1 y_2)}$$

$$= p(y_2|x_2)\,p(x_2|y_1)\frac{p(y_1)}{\cancel{p(x_2)}}\cancel{p(x_2)}\frac{1}{p(y_1 y_2)}$$

$$\propto p(y_2|x_2)\,p(x_2|y_1)\frac{1}{p(y_2|y_1)}$$

$$p(y_2|x_2) = N(Hx_2, R) \quad \text{we know it!}$$

$$p(x_2|y_1) = ?$$

$$x_2 = M(x_1)$$

$$\Rightarrow x_2|y_1 = M(\underbrace{x_1|y_1}_{N(\mu_1, P_1)})$$

$$x_2|y_1 = N(\underbrace{M\mu_1}_{\text{forecast mean.}}, \underbrace{MP_1M^T}_{\text{Forecast Covariance}})$$

$\Rightarrow$ Second step:
$$\mu_f = M\mu_1 \qquad P_f = MP_1M^T \rightarrow x_2 = \mu_f + K(y_2 - H\mu_f)$$
$$P_2 = (I - KH)P_f.$$
with $K = P_f H^T(HP_fH^T + R)^{-1}$

In general: recursion.

Given $x_k \mid y_{1:k} \sim \mathcal{N}(\mu_k, P_k)$

Forecast: $\mu_f = M \mu_k$

$$P_f = M P_k M^T$$

$$K = P_f H^T (H P_f H^T + R)^{-1}$$

$$\mu_{k+1} = \mu_f + K(y_{k+1} - H\mu_f)$$

$$P_{k+1} = (I - KH) P_k$$

---

A closer look at the covariance matrix

$$P_{k+1} = (I - KH) P_f$$

$$P_f = M P_k M^T$$

$$K = P_f H^T (H P_f H^T + R)^{-1}$$

Assume steady state: $P_k = P_{k+1} = P$

$K$ is const.

Call $P_f = M P M^T = X$

$$P = (I - KH)X$$

$$X = MPM^T$$

$$K = XH^T(HPH^T + R)^{-1}$$

Combine

$$P = X - XH^T(HPH^T + R)^{-1}HX$$

$$\underbrace{MPM^T}_{X} = MXM^T - MXH^T(HPH^T + R)^{-1}HXM$$

$$\underbrace{X = MXM^T - MXH^T(HPH^T + R)^{-1}HXM}_{\text{Riccati Eqn.}}$$

Algebraic Riccati eqns, P. Lancaster Leiba Rodman.

→ Kalman gain & Posterior covariance quickly reach steady state if M, R, H are constants. You will see this in HW.