The applet visualizes the first two steps of a simple Kalman filter. We follow the notations and terminology from the Welch & Bishop’s paper [1].

1. Estimating a random constant

Let $x$ and $v_1, v_2$ be independent centered random variables with finite 2nd moments:

\[ P_0 = E(x^2); \]
\[ R = E(v_1)^2 = E(v_2)^2. \]

Let $z_1 = x + v_1$ and $z_2 = x + v_2$.

We interpret $z_1$ and $z_2$ as noisy measurements of $x$.

**Problem.** Given $z_1$ and $z_2$, find the best linear estimate $\hat{x} = c_1 z_1 + c_2 z_2$ of $x$ in the sense that $E((x - \hat{x})^2)$ is as small as possible.

Moreover, perform this task recursively:
1. First, approximate $x$ by a multiple of $z_1$; that would produce an estimate $\hat{x}_1 = K z_1$ for some $K$;
2. Then, determine $\hat{x}$ as a linear combination of $\hat{x}_1$ and $z_2$.

2. Geometric Interpretation of the problem

- The space $\text{Span}(x, z_1, z_2) = \{ c_0 x + c_1 z_1 + c_2 z_2 | c_i \in \mathbb{R} \}$ with the scalar product $u \cdot v = E(uv)$ is a Euclidean space.
- Consider the norm $||u|| = \sqrt{u \cdot u}$. In our notations:
  \[ ||x||^2 = P_0; \]
  \[ ||v_1||^2 = ||v_2||^2 = R. \]

**Problem.** Find the orthogonal projection $\hat{x}_2$ of the variable $x$ onto the subspace $\text{Span}(z_1, z_2) = \{ c_1 z_1 + c_2 z_2 | c_i \in \mathbb{R} \}$.

Specifically, implement this as a recursive procedure:
1. First find the projection $\hat{x}_1$ of $x$ onto $\text{Span}(z_1)$;
2. Then, compute $\hat{x}_2$ as a linear combination of $\hat{x}_1$ and $z_2$.

3. Kalman estimates

- If $\hat{x}_1 = K_1 z_1$ is the orthogonal projection of $x$ onto the vector $z_1$, then
  \[ K_1 = \frac{x \cdot z_1}{||z_1||^2} = \frac{x \cdot (x + v_1)}{||x + v_1||^2} = \frac{||x||^2 + x \cdot v_1}{||x||^2 + ||v_1||^2} = \frac{P_0}{P_0 + R} \]
  since $x$ and $v_1$ are orthogonal.
- The variance of $\hat{x}_1$ is
  \[ ||\hat{x}_1||^2 = ||K_1 z_1||^2 = (K_1)^2 (P_0 + R)^2 = K_1 P_0. \]
- Now, what Welch and Bishop[1] denote by $P_1$ is
  \[ ||x - \hat{x}_1||^2 = ||x||^2 - ||\hat{x}_1||^2 = P_0 - K_1 P_0 = (1 - K_1) P_0 \]
  by the Pythagorean theorem.
• \( z_2 - \hat{x}_1 \) and \( z_1 \) are orthogonal since their dot product is

\[
(z_2 - \hat{x}_1) \cdot z_1 = (v_2 + x - \hat{x}_1) \cdot (v_1 + x) = v_2 \cdot (v_1 + x) + (x - \hat{x}_1) \cdot v_1 + (x - \hat{x}_1) \cdot x = 0 + 0 + 0.
\]

Thus, \( z_2 - \hat{x}_1 \) and \( z_1 \) form an orthogonal basis in the plane spanned by \( z_1 \) and \( z_2 \).

• Therefore, the projection of \( x \) on \( \text{span}(z_1, z_2) \) is the sum of the projections on \( z_1 \) and \( z_2 - \hat{x}_1 \):

\[
\hat{x}_2 = \hat{x}_1 + K_2 (z_2 - \hat{x}_1),
\]

where

\[
K_2 = \frac{x \cdot (z_2 - \hat{x}_1)}{||z_2 - \hat{x}_1||^2} = \frac{(x - \hat{x}_1) \cdot (z_2 - \hat{x}_1)}{||x - \hat{x}_1 + v_2||^2} = \frac{||x - \hat{x}_1||^2}{||x - \hat{x}_1||^2 + R} = \frac{P_1}{P_1 + R}.
\]

• The variance of \( K_2 (z_2 - \hat{x}_1) \) is

\[
||K_2(z_2 - \hat{x}_1)||^2 = (K_2)^2 ||z_2 - \hat{x}_1||^2 = K_2 \frac{P_1}{(P_1 + R)(P_1 + R)}.
\]

• What Welch and Bishop would denote by \( P_2 \) is

\[
||x - \hat{x}_2||^2 = ||x - \hat{x}_1||^2 - ||K_2(z_2 - \hat{x}_1)||^2 = P_1 - K_2 P_1 = (1 - K_2) P_1
\]

by the Pythagorean theorem. See the illustration.

References