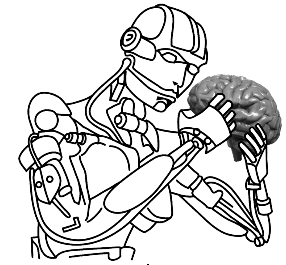
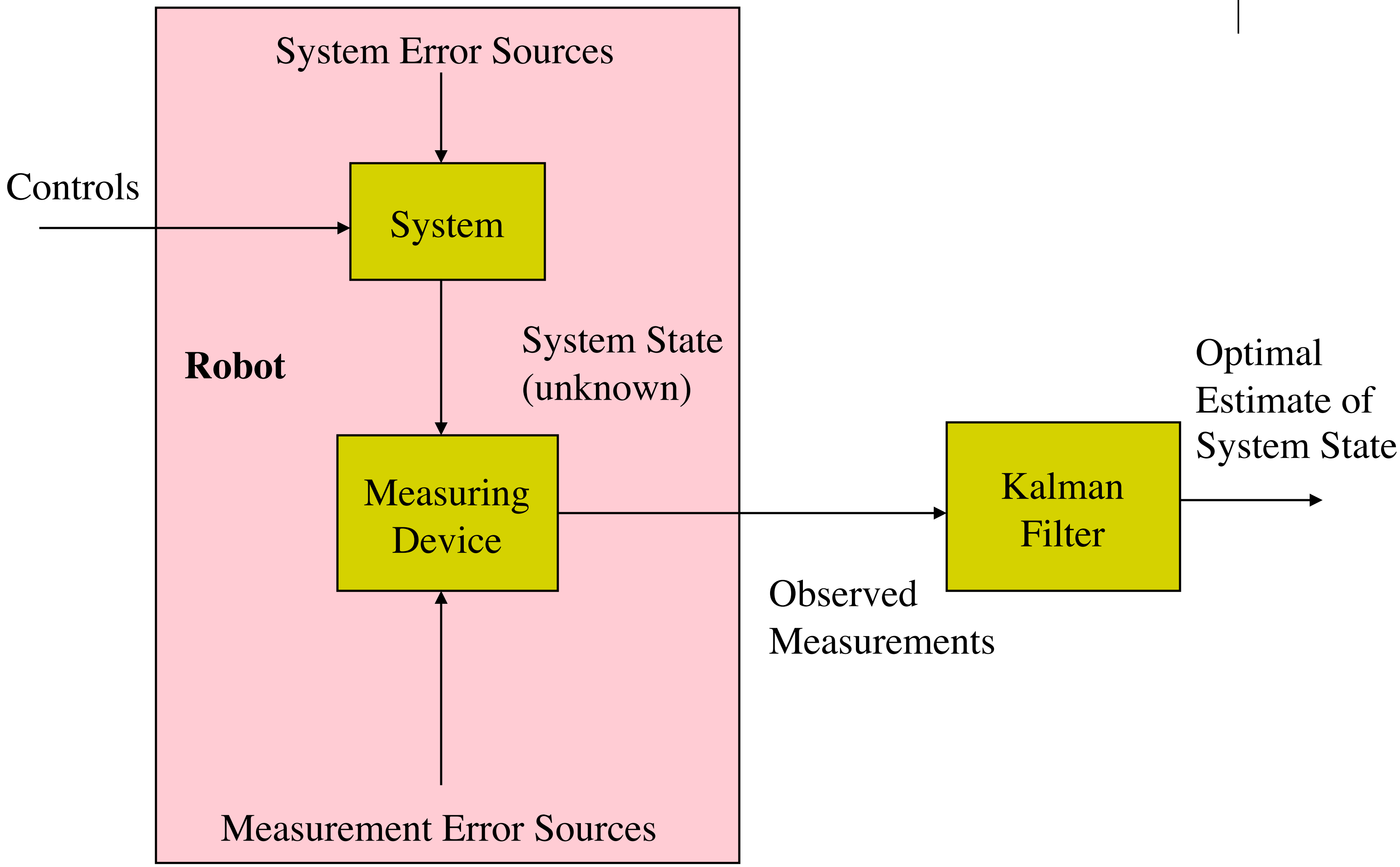
CS545—Contents XVIII

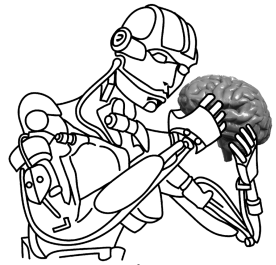
- Kalman Filtering
  - The Kalman filter framework
  - Derivation of Kalman filter update equations
- Reading Assignment for Next Class
  - See http://www-clmc.usc.edu/~cs545

A Typical Kalman Filter Application



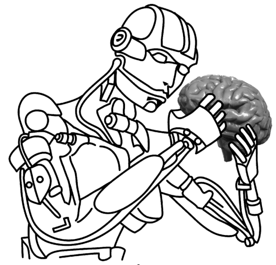


When to use a Kalman Filter



- Eliminate noise in measurements
- Generate non-observable states (e.g., velocities from position signals)
- For prediction of future states (systems with time delays)
- Optimal filtering

The Kalman Filter Framework (1960)



- Given:
  - A discrete stochastic linear controlled dynamical system

$$\mathbf{x}^{n+1} = \mathbf{A}\mathbf{x}^n + \mathbf{B}\mathbf{u}^n + \mathbf{w}^n$$

Model uncertainty

  - A measurement function

$$\mathbf{y}^n = \mathbf{C}\mathbf{x}^n + \mathbf{v}^n$$

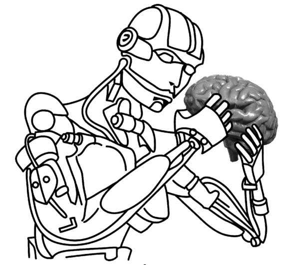
Measurement noise

  - Some knowledge about the additive noise

$$E\{\mathbf{w}\} = 0,\ E\{\mathbf{v}\} = 0,\ E\left\{\mathbf{w}\mathbf{w}^T\right\} = \mathbf{Q},\ E\left\{\mathbf{v}\mathbf{v}^T\right\} = \mathbf{R},\ E\left\{\mathbf{w}\mathbf{v}^T\right\} = 0$$

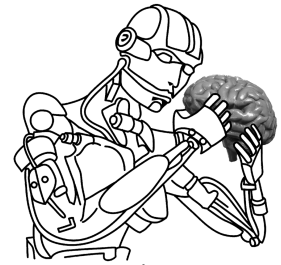
- Goal:
  - Find the best (recursive) estimate of the state **x** of the system.

Properties of the Kalman Filter



- Allows to estimate past, present, and future states
- Requires a model of the system dynamics (at least approximate)
- Much better than digital filters (why?)
- The Kalman filter is optimal for linear systems
- Extensions to nonlinear systems exist: "Extended Kalman Filter"
- The Kalman Filter can be calculated in the same way as gains for Linear Quadratic Regulator problems.

In Which Sense is the Kalman Filter Optimal?



- Assume an *a priori* estimate of the state **x** at step *n* given the knowledge of the process dynamics and the previous state estimate at *n-1*:

$$\tilde{\mathbf{x}}^n = \mathbf{A}\hat{\mathbf{x}}^{n-1} + \mathbf{B}\mathbf{u}^{n-1}$$

- Additionally we measure the output of the process at *n:*

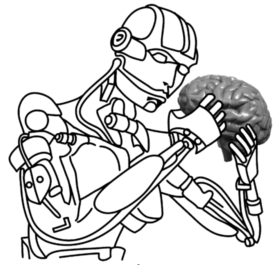
$$\mathbf{y}^n$$

- How can we optimally (linearly) combine the estimate and measurement to obtain the best reconstruction of the true **x**?

$$\hat{\mathbf{x}}^n = \tilde{\mathbf{x}}^n + \mathbf{K}\left(\mathbf{y}^n - \mathbf{C}\tilde{\mathbf{x}}^n\right)$$

This is (one form) of the famous Kalman update equation!

In Which Sense is the Kalman Filter Optimal? (cont'd)



- The **K** matrix is the open parameter in the Kalman filter.
- We want to choose **K** such that we minimize the *a posteriori* estimation error (in expectation):

$$\mathbf{e}^n = \mathbf{x}^n - \hat{\mathbf{x}}^n$$

  - I.e., minimize the expected error covariance

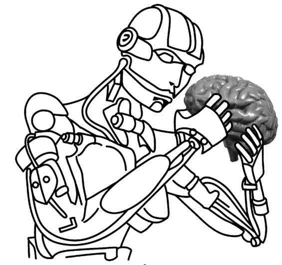
$$E\left\{\mathbf{e}^n \mathbf{e}^{nT}\right\}$$

  - The Kalman filter gains are derived by minimizing the posterior error covariance, resulting in:

$$\mathbf{K}^n = \tilde{\mathbf{P}}^n \mathbf{C}^T \left(\mathbf{C}^n \tilde{\mathbf{P}}^n \mathbf{C}^T + \mathbf{R}\right)^{-1}$$

with: $\tilde{\mathbf{e}}^n = \mathbf{x}^n - \tilde{\mathbf{x}}^n$ and $\tilde{\mathbf{P}}^n = E\left\{\tilde{\mathbf{e}}^n \tilde{\mathbf{e}}^{nT}\right\}$ (prior error covariance)

Derivation of Kalman Gains



$$E\left\{\mathbf{e}^n\mathbf{e}^{nT}\right\} = E\left\{\left(\mathbf{x}^n - \hat{\mathbf{x}}^n\right)\left(\mathbf{x}^n - \hat{\mathbf{x}}^n\right)^T\right\}$$

$$= E\left\{\left(\mathbf{x}^n - \tilde{\mathbf{x}}^n - \mathbf{K}\left(\mathbf{y}^n - \mathbf{C}\tilde{\mathbf{x}}^n\right)\right)\left(\mathbf{x}^n - \tilde{\mathbf{x}}^n - \mathbf{K}\left(\mathbf{y}^n - \mathbf{C}\tilde{\mathbf{x}}^n\right)\right)^T\right\}$$

$$= E\left\{\left(\mathbf{x}^n - \tilde{\mathbf{x}}^n\right)\left(\mathbf{x}^n - \tilde{\mathbf{x}}^n\right)^T\right\} + E\left\{\left(\mathbf{K}\left(\mathbf{y}^n - \mathbf{C}\tilde{\mathbf{x}}^n\right)\right)\left(\mathbf{K}\left(\mathbf{y}^n - \mathbf{C}\tilde{\mathbf{x}}^n\right)\right)^T\right\} -$$

$$E\left\{\left(\mathbf{x}^n - \tilde{\mathbf{x}}^n\right)\left(\mathbf{K}\left(\mathbf{y}^n - \mathbf{C}\tilde{\mathbf{x}}^n\right)\right)^T\right\} - E\left\{\left(\mathbf{K}\left(\mathbf{y}^n - \mathbf{C}\tilde{\mathbf{x}}^n\right)\right)\left(\mathbf{x}^n - \tilde{\mathbf{x}}^n\right)^T\right\}$$

$$= \tilde{\mathbf{P}}^n + \mathbf{K}\,E\left\{\left(\mathbf{y}^n - \mathbf{C}\tilde{\mathbf{x}}^n\right)\left(\mathbf{y}^n - \mathbf{C}\tilde{\mathbf{x}}^n\right)^T\right\}\mathbf{K}^T - E\left\{\left(\mathbf{x}^n - \tilde{\mathbf{x}}^n\right)\left(\mathbf{y}^n - \mathbf{C}\tilde{\mathbf{x}}^n\right)^T\right\}\mathbf{K}^T - \mathbf{K}\,E\left\{\left(\mathbf{y}^n - \mathbf{C}\tilde{\mathbf{x}}^n\right)\left(\mathbf{x}^n - \tilde{\mathbf{x}}^n\right)^T\right\}$$

$$= \tilde{\mathbf{P}}^n + \mathbf{K}\,E\left\{\left(\mathbf{C}\mathbf{x}^n + \mathbf{v}^n - \mathbf{C}\tilde{\mathbf{x}}^n\right)\left(\mathbf{C}\mathbf{x}^n + \mathbf{v}^n - \mathbf{C}\tilde{\mathbf{x}}^n\right)^T\right\}\mathbf{K}^T - E\left\{\left(\mathbf{x}^n - \tilde{\mathbf{x}}^n\right)\left(\mathbf{C}\mathbf{x}^n + \mathbf{v}^n - \mathbf{C}\tilde{\mathbf{x}}^n\right)^T\right\}\mathbf{K}^T$$

$$-\mathbf{K}\,E\left\{\left(\mathbf{C}\mathbf{x}^n + \mathbf{v}^n - \mathbf{C}\tilde{\mathbf{x}}^n\right)\left(\mathbf{x}^n - \tilde{\mathbf{x}}^n\right)^T\right\}$$

$$= \tilde{\mathbf{P}}^n + \mathbf{K}\mathbf{C}\,E\left\{\left(\mathbf{x}^n - \tilde{\mathbf{x}}^n\right)\left(\mathbf{x}^n - \tilde{\mathbf{x}}^n\right)^T\right\}\mathbf{C}^T\mathbf{K}^T + \mathbf{K}\mathbf{R}\mathbf{K}^T - E\left\{\left(\mathbf{x}^n - \tilde{\mathbf{x}}^n\right)\left(\mathbf{x}^n - \tilde{\mathbf{x}}^n\right)^T\right\}\mathbf{C}^T\mathbf{K}^T - \mathbf{K}\mathbf{C}E\left\{\left(\mathbf{x}^n - \tilde{\mathbf{x}}^n\right)\left(\mathbf{x}^n - \tilde{\mathbf{x}}^n\right)^T\right\}$$

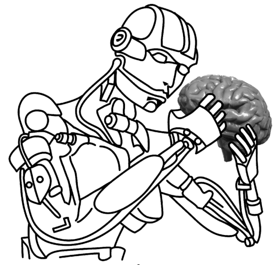
$$= \tilde{\mathbf{P}}^n + \mathbf{K}\mathbf{C}\tilde{\mathbf{P}}^n\mathbf{C}^T\mathbf{K}^T + \mathbf{K}\mathbf{R}\mathbf{K}^T - \tilde{\mathbf{P}}^n\mathbf{C}^T\mathbf{K}^T - \mathbf{K}\mathbf{C}\tilde{\mathbf{P}}^n$$

$$\frac{\partial E\left\{\mathbf{e}^n\mathbf{e}^{nT}\right\}}{\partial \mathbf{K}} = 2\mathbf{C}\tilde{\mathbf{P}}^n\mathbf{C}^T\mathbf{K}^T + 2\mathbf{R}\mathbf{K}^T - \mathbf{C}\tilde{\mathbf{P}}^n - \left(\tilde{\mathbf{P}}^n\mathbf{C}^T\right)^T = 0$$

$$\mathbf{K}^T = \left(\mathbf{C}\,\tilde{\mathbf{P}}^n\mathbf{C}^T + \mathbf{R}\right)^{-1}\mathbf{C}\tilde{\mathbf{P}}^n$$

$$\mathbf{K} = \tilde{\mathbf{P}}^n\mathbf{C}^T\left(\mathbf{C}\,\tilde{\mathbf{P}}^n\mathbf{C}^T + \mathbf{R}\right)^{-1}$$

Derivation of Posterior Covariance Update



$$\mathbf{P}^n = E\left\{\mathbf{e}^n \mathbf{e}^{n^T}\right\} =$$

$$= \tilde{\mathbf{P}}^n + \mathbf{K}\mathbf{C}\tilde{\mathbf{P}}^n\mathbf{C}^T\mathbf{K}^T + \mathbf{K}\mathbf{R}\mathbf{K}^T - \tilde{\mathbf{P}}^n\mathbf{C}^T\mathbf{K}^T - \mathbf{K}\mathbf{C}\tilde{\mathbf{P}}^n$$

$$= \tilde{\mathbf{P}}^n + \mathbf{K}\left(\mathbf{C}\tilde{\mathbf{P}}^n\mathbf{C}^T + \mathbf{R}\right)\mathbf{K}^T - \tilde{\mathbf{P}}^n\mathbf{C}^T\mathbf{K}^T - \mathbf{K}\mathbf{C}\tilde{\mathbf{P}}^n$$

$$= \tilde{\mathbf{P}}^n + \tilde{\mathbf{P}}^n\mathbf{C}^T\left(\mathbf{C}\tilde{\mathbf{P}}^n\mathbf{C}^T + \mathbf{R}\right)^{-1}\left(\mathbf{C}\tilde{\mathbf{P}}^n\mathbf{C}^T + \mathbf{R}\right)\left(\mathbf{C}\tilde{\mathbf{P}}^n\mathbf{C}^T + \mathbf{R}\right)^{-1}\mathbf{C}\tilde{\mathbf{P}}^n$$

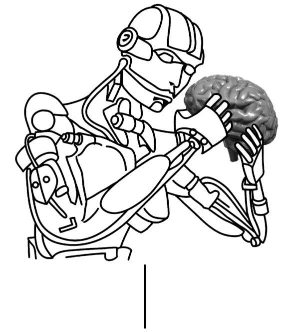
$$-\tilde{\mathbf{P}}^n\mathbf{C}^T\left(\mathbf{C}\tilde{\mathbf{P}}^n\mathbf{C}^T + \mathbf{R}\right)^{-1}\mathbf{C}\tilde{\mathbf{P}}^n - \mathbf{K}\mathbf{C}\tilde{\mathbf{P}}^n$$

$$= \tilde{\mathbf{P}}^n + \tilde{\mathbf{P}}^n\mathbf{C}^T\left(\mathbf{C}\tilde{\mathbf{P}}^n\mathbf{C}^T + \mathbf{R}\right)^{-1}\mathbf{C}\tilde{\mathbf{P}}^n - \tilde{\mathbf{P}}^n\mathbf{C}^T\left(\mathbf{C}\tilde{\mathbf{P}}^n\mathbf{C}^T + \mathbf{R}\right)^{-1}\mathbf{C}\tilde{\mathbf{P}}^n - \mathbf{K}\mathbf{C}\tilde{\mathbf{P}}^n$$

$$= \tilde{\mathbf{P}}^n - \mathbf{K}\mathbf{C}\tilde{\mathbf{P}}^n$$

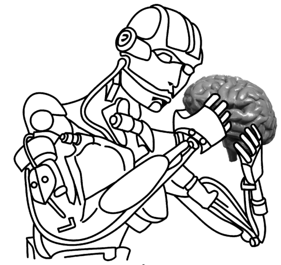
$$= (\mathbf{I} - \mathbf{K}\mathbf{C})\tilde{\mathbf{P}}^n$$

Derivation of Prior Covariance at the Next Step



$$
\begin{aligned}
\tilde{\mathbf{P}}^{n+1} &= E\left\{\tilde{\mathbf{e}}^{n+1}\tilde{\mathbf{e}}^{n+1\,T}\right\} = E\left\{\left(\mathbf{x}^{n+1} - \tilde{\mathbf{x}}^{n+1}\right)\left(\mathbf{x}^{n+1} - \tilde{\mathbf{x}}^{n+1}\right)^{T}\right\} \\
&= E\left\{\left(\mathbf{A}\mathbf{x}^{n} + \mathbf{B}\mathbf{u}^{n} + \mathbf{w}^{n} - \mathbf{A}\hat{\mathbf{x}}^{n} - \mathbf{B}\mathbf{u}^{n}\right)\left(\mathbf{A}\mathbf{x}^{n} + \mathbf{B}\mathbf{u}^{n} + \mathbf{w}^{n} - \mathbf{A}\hat{\mathbf{x}}^{n} - \mathbf{B}\mathbf{u}^{n}\right)^{T}\right\} \\
&= E\left\{\left(\mathbf{A}\mathbf{x}^{n} + \mathbf{w}^{n} - \mathbf{A}\hat{\mathbf{x}}^{n}\right)\left(\mathbf{A}\mathbf{x}^{n} + \mathbf{w}^{n} - \mathbf{A}\hat{\mathbf{x}}^{n}\right)^{T}\right\} \\
&= E\left\{\left(\mathbf{A}\mathbf{x}^{n} - \mathbf{A}\hat{\mathbf{x}}^{n}\right)\left(\mathbf{A}\mathbf{x}^{n} - \mathbf{A}\hat{\mathbf{x}}^{n}\right)^{T} + \mathbf{w}^{n}\left(\mathbf{A}\mathbf{x}^{n} - \mathbf{A}\hat{\mathbf{x}}^{n}\right)^{T} + \left(\mathbf{A}\mathbf{x}^{n} - \mathbf{A}\hat{\mathbf{x}}^{n}\right)\mathbf{w}^{n\,T} + \mathbf{w}^{n}\mathbf{w}^{n\,T}\right\} \\
&= \mathbf{A}E\left\{\left(\mathbf{x}^{n} - \hat{\mathbf{x}}^{n}\right)\left(\mathbf{x}^{n} - \hat{\mathbf{x}}^{n}\right)\right\}\mathbf{A}^{T} + E\left\{\mathbf{w}^{n}\mathbf{w}^{n\,T}\right\} \\
&= \mathbf{A}\mathbf{P}^{n}\mathbf{A}^{T} + \mathbf{Q}
\end{aligned}
$$

Summary: The Discrete Kalman Filter Equations



- The time update equations:

$$\tilde{\mathbf{x}}^n = \mathbf{A}\hat{\mathbf{x}}^{n-1} + \mathbf{B}\mathbf{u}^{n-1}$$

$$\tilde{\mathbf{P}}^{n+1} = \mathbf{A}^n \mathbf{P}^n \mathbf{A}^{n^T} + \mathbf{Q}$$

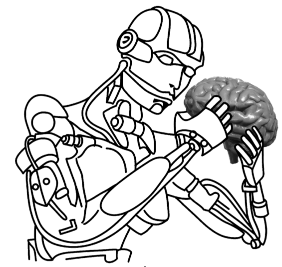
- The measurement update equations

$$\mathbf{K}^n = \tilde{\mathbf{P}}^n \mathbf{C}^T \left(\mathbf{C}^n \tilde{\mathbf{P}}^n \mathbf{C}^T + \mathbf{R} \right)^{-1}$$

$$\hat{\mathbf{x}}^n = \tilde{\mathbf{x}}^n + \mathbf{K}\left(\mathbf{y}^n - \mathbf{C}\tilde{\mathbf{x}}^n \right)$$

$$\mathbf{P}^n = \left(\mathbf{I} - \mathbf{K}^n \mathbf{C} \right) \tilde{\mathbf{P}}^n$$

Discussion



- What happens if the a priori estimate of the process noise is zero?

$$\mathbf{K} = 0$$

- What happens if the measurement noise is zero?

$$\mathbf{K} = \mathbf{C}^{-1}$$

Example

- Estimate a constant from noisy data:

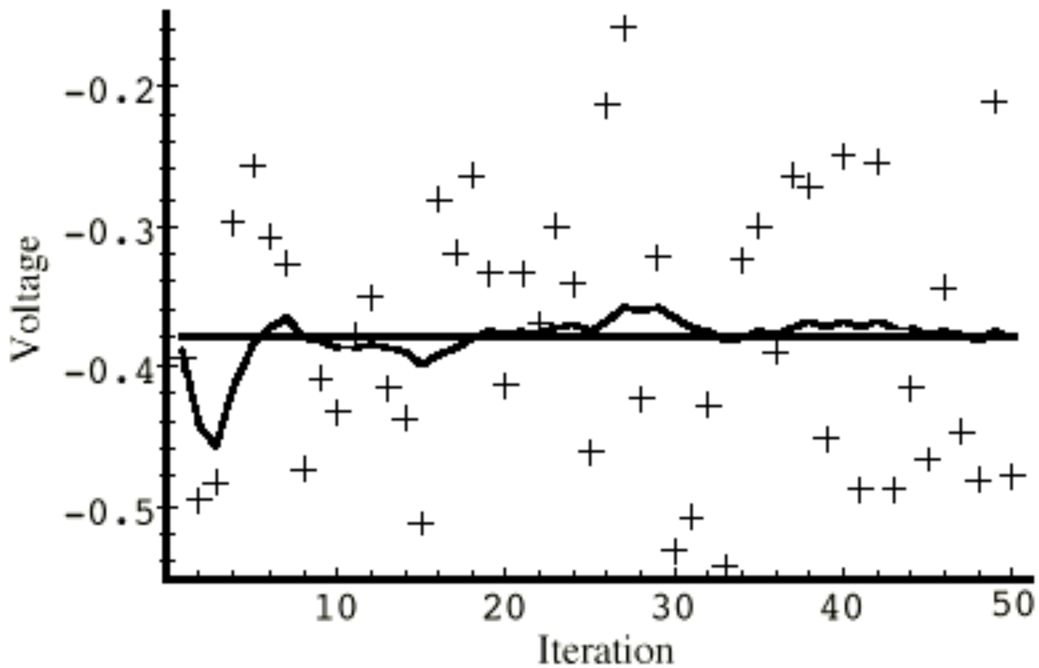


Figure 3-1. The first simulation: $R = (0.1)^2 = 0.01$. The true value of the random constant $x = -0.37727$ is given by the solid line, the noisy measurements by the cross marks, and the filter estimate by the remaining curve.

Example (cont'd)

- Estimate a constant from noisy data:

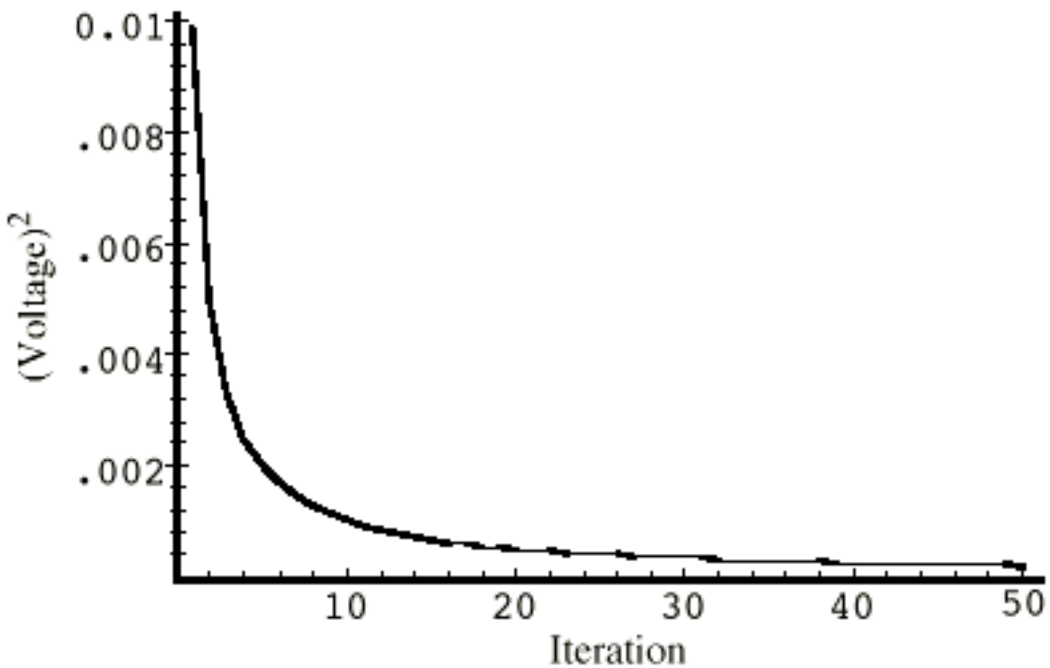


Figure 3-2. After 50 iterations, our initial (rough) error covariance P_k^- choice of 1 has settled to about 0.0002 (Volts^2).