Lecture 8
The Extended Kalman filter

- Nonlinear filtering
- Extended Kalman filter
- Linearization and random variables

Nonlinear filtering

- nonlinear Markov model:

$$x(t+1) = f(x(t), w(t)), \qquad y(t) = g(x(t), v(t))$$

  - f is (possibly nonlinear) dynamics function
  - g is (possibly nonlinear) measurement or output function
  - $w(0), w(1), \ldots, v(0), v(1), \ldots$ are independent
  - even if w, v Gaussian, x and y need not be

- nonlinear filtering problem: find, $e.g.$,

$$\hat{x}(t|t-1) = \mathbf{E}(x(t)|y(0), \ldots, y(t-1)), \qquad \hat{x}(t|t) = \mathbf{E}(x(t)|y(0), \ldots, y(t))$$

- general nonlinear filtering solution involves a PDE, and is not practical

Extended Kalman filter

- extended Kalman filter (EKF) is *heuristic* for nonlinear filtering problem
- often works well (when tuned properly), but sometimes not
- widely used in practice
- based on
  - linearizing dynamics and output functions at current estimate
  - propagating an approximation of the conditional expectation and covariance

Linearization and random variables

- consider $\phi : \mathbf{R}^n \to \mathbf{R}^m$
- suppose $\mathbf{E}\, x = \bar{x}$, $\mathbf{E}(x - \bar{x})(x - \bar{x})^T = \Sigma_x$, and $y = \phi(x)$
- if Σ_x is small, ϕ is not too nonlinear,

$$y \approx \tilde{y} = \phi(\bar{x}) + D\phi(\bar{x})(x - \bar{x})$$

$$\tilde{y} \sim \mathcal{N}(\phi(\bar{x}), D\phi(\bar{x})\Sigma_x D\phi(\bar{x})^T)$$

- gives *approximation* for mean and covariance of nonlinear function of random variable:

$$\bar{y} \approx \phi(\bar{x}), \qquad \Sigma_y \approx D\phi(\bar{x})\Sigma_x D\phi(\bar{x})^T$$

- if Σ_x is not small compared to 'curvature' of ϕ, these estimates are poor

- a good estimate can be found by Monte Carlo simulation:

$$\bar{y} \approx \bar{y}^{\mathrm{mc}} = \frac{1}{N}\sum_{i=1}^{N} \phi(x^{(i)})$$

$$\Sigma_y \approx \frac{1}{N}\sum_{i=1}^{N} \left(\phi(x^{(i)}) - \bar{y}^{\mathrm{mc}}\right)\left(\phi(x^{(i)}) - \bar{y}^{\mathrm{mc}}\right)^T$$

  where $x^{(1)}, \ldots, x^{(N)}$ are samples from the distribution of x, and N is large

- another method: use Monte Carlo formulas, with a small number of nonrandom samples chosen as 'typical', *e.g.*, the 90% confidence ellipsoid semi-axis endpoints

$$x^{(i)} = \bar{x} \pm \beta v_i, \qquad \Sigma_x = V\Lambda V^T$$

Example

$x \sim \mathcal{N}(0,1)$, $y = \exp(x)$

(for this case we can compute mean and variance of y exactly)

	$\bar{y}$	σ_y
exact values	$e^{1/2} = 1.649$	$\sqrt{e^2 - e} = 2.161$
linearization	1.000	1.000
Monte Carlo ($N = 10$)	1.385	1.068
Monte Carlo ($N = 100$)	1.430	1.776
Sigma points ($x = \bar{x},\ \bar{x} \pm 1.5\sigma_x$)	1.902	2.268

Extended Kalman filter

- *initialization:* $\hat{x}(0|-1) = \bar{x}_0$, $\Sigma(0|-1) = \Sigma_0$
- *measurement update*
  - linearize output function at $x = \hat{x}(t|t-1)$:

$$
\begin{aligned}
C &= \frac{\partial g}{\partial x}(\hat{x}(t|t-1), 0) \\
V &= \frac{\partial g}{\partial v}(\hat{x}(t|t-1), 0)\Sigma_v \frac{\partial g}{\partial v}(\hat{x}(t|t-1), 0)^T
\end{aligned}
$$

  - measurement update based on linearization

$$
\begin{aligned}
\hat{x}(t|t) &= \hat{x}(t|t-1) + \Sigma_{t|t-1} C^T \left(C\Sigma_{t|t-1}C^T + V\right)^{-1} \ldots \\
&\quad \ldots (y(t) - g(\hat{x}(t|t-1), 0)) \\
\Sigma_{t|t} &= \Sigma_{t|t-1} - \Sigma_{t|t-1} C^T \left(C\Sigma_{t|t-1}C^T + V\right)^{-1} C\Sigma_{t|t-1}
\end{aligned}
$$

- *time update*
  - linearize dynamics function at $x = \hat{x}(t|t)$:

$$
\begin{aligned}
A &= \frac{\partial f}{\partial x}(\hat{x}(t|t), 0) \\
W &= \frac{\partial f}{\partial w}(\hat{x}(t|t), 0)\Sigma_w \frac{\partial f}{\partial w}(\hat{x}(t|t), 0)^T
\end{aligned}
$$

  - time update based on linearization

$$
\hat{x}(t+1|t) = f(\hat{x}(t|t), 0), \qquad \Sigma_{t+1|t} = A\Sigma_{t|t}A^T + W
$$

- replacing linearization with Monte Carlo yields *particle filter*
- replacing linearization with sigma-point estimates yields *unscented Kalman filter* (UKF)

Example

- $p(t)$, $u(t) \in \mathbf{R}^2$ are position and velocity of vehicle, with $(p(0), u(0)) \sim \mathcal{N}(0, I)$
- vehicle dynamics:

$$p(t+1) = p(t) + 0.1u(t), \qquad u(t+1) = \begin{bmatrix} 0.85 & 0.15 \\ -0.1 & 0.85 \end{bmatrix} u(t) + w(t)$$

  $w(t)$ are IID $\mathcal{N}(0, I)$

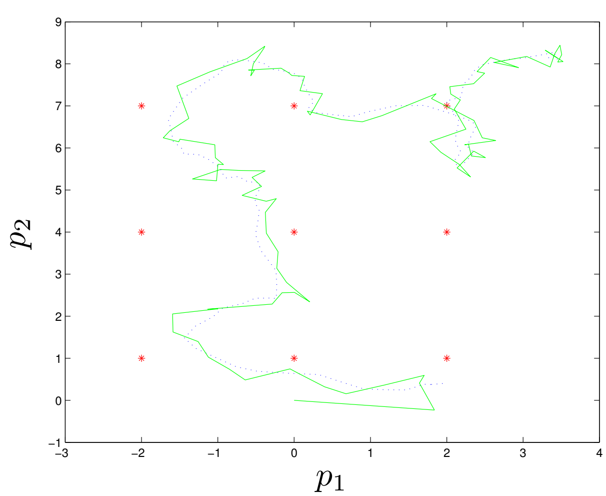
- measurements: noisy measurements of distance to 9 points $p_i \in \mathbf{R}^2$

$$y_i(t) = \|p(t) - p_i\| + v_i(t), \quad i = 1, \ldots, 9,$$

  $v_i(t)$ are IID $\mathcal{N}(0, 0.3^2)$

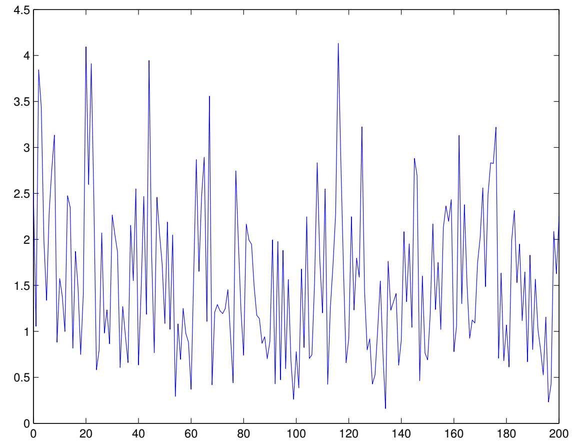
EKF results

- EKF initialized with $\hat{x}(0|-1) = 0$, $\Sigma(0|-1) = I$, where $x = (p, u)$
- p_i shown as stars; $p(t)$ as dotted curve; $\hat{p}(t|t)$ as solid curve



Current position estimation error

$\|\hat{p}(t|t) - p(t)\|$ versus t



Current position estimation predicted error

$(\Sigma(t|t)_{11} + \Sigma(t|t)_{22})^{1/2}$ versus t