# Guidance, Navigation and Control ELEC3224

Professor Daniel E. Clark

ECS

*daniel.clark@soton.ac.uk*

# Multivariate Gaussian

Suppose that we know that these measurements are distributed according to a multivariate Gaussian likelihood function $g(z \mid \theta)$, where our parameter $\theta$ is now 2-dimensional, i.e.

$$g(z \mid \theta) = \frac{1}{(2\pi)\sqrt{|\det(R)|}} \exp\left(-\frac{1}{2}(z-\theta)^T R^{-1}(z-\theta)\right),$$

where $R$ is the $2 \times 2$ covariance matrix (the 2 -dimensional version of the variance), $\theta$ is the mean parameter, and $z$ is the observation vector.

# Bayesian estimation

If we specify the prior $p(\theta)$ to be a Gaussian, i.e.

$$p(\theta) = \frac{1}{(2\pi)\sqrt{|\det(P)|}} \exp\left(-\frac{1}{2}(\theta - m)^T R^{-1}(\theta - m)\right),$$

where the mean and covariance of the Gaussian are given by $m$ and $P$ respectively, then the posterior distribution updated with measurement $z$ also happens to be a Gaussian, i.e.

$$p(\theta \mid z) = \frac{1}{(2\pi)\sqrt{|\det(\bar{P})|}} \exp\left(-\frac{1}{2}(\theta - \bar{m})^T \bar{P}^{-1}(\theta - \bar{m})\right),$$

where $\bar{m}$ and $\bar{P}$ are the new mean and covariance.

# Single-object Bayesian filtering

In many dynamic state estimation problems, the state is assumed to follow a Markov process on the state space $\mathcal{X} \subseteq R^{n_x}$, with transition density $f_{k|k-1}(\cdot \mid \cdot)$, i.e. given a state $x_{k-1}$ at time $k-1$, the probability density of a transition to the state $x_k$ at time $k$ is

$$f_{k|k-1}\left(x_k \mid x_{k-1}\right)$$

# Single-object Bayesian filtering

The Markov process is partially observed in the observation space $\mathcal{Z} \subseteq R^{n_z}$, as modelled by the likelihood function $g_k(\cdot \mid \cdot)$, i.e. given a state $x_k$ at time $k$, the probability density of receiving the observation $z_k \in \mathcal{Z}$ is

$$g_k\left(z_k \mid x_k\right)$$

# Single-object Bayesian filtering

All information about the state at time $k$ is encapsulated in the posterior (or filtering ) density

$$p_k\left(\cdot \mid z_{1:k}\right)$$

which is defined as the probability density of the state at time $k$ given the measurement history $z_{1:k}$ up to time $k$. From an initial density $p_0(\cdot)$, the posterior density at time $k$ can be computed using the Bayes recursion

$$p_{k|k-1}\left(x_k \mid z_{1:k-1}\right) = \int f_{k|k-1}\left(x_k \mid x\right) p_{k-1}\left(x \mid z_{1:k-1}\right) dx,$$

$$p_k\left(x_k \mid z_{1:k}\right) = \frac{g_k\left(z_k \mid x_k\right) p_{k|k-1}\left(x_k \mid z_{1:k-1}\right)}{\int g_k\left(z_k \mid x\right) p_{k|k-1}\left(x \mid z_{1:k-1}\right) dx}.$$

# The Kalman filter

The Kalman filter provides an optimal solution to the Bayes' filter under the assumptions that the dynamics and observation characteristics of the object is linear, with Gaussian noise, so that

$$f_{k|k-1}(x \mid y) = \mathcal{N}\left(x; F_{k-1}y, Q_{k-1}\right)$$
$$g_k(z \mid x) = \mathcal{N}\left(z; H_k x, R_k\right),$$

where $\mathcal{N}(z; m, P)$ denotes a multi-variate Gaussian density with mean $m$ and covariance $P$, i.e.

$$\mathcal{N}(z; m, P) = |2\pi P|^{-1/2} \exp\left(-\frac{1}{2}(x-m)^T P^{-1}(x-m)\right)$$

$F_{k-1}$ is a state transition matrix, $Q_{k-1}$ is the process noise covariance, $H_k$ is the matrix that projects the state onto the observation space, and $R_k$ is the observation covariance noise matrix.

# The Kalman filter

The Kalman filter makes use of the following two identities for Gaussians Identity 1: Given $F, Q, m, P$ of suitable dimensions, and $Q$ and $P$ are positive definite, then

$$\int \mathcal{N}(x; Fy, Q)\mathcal{N}(y; m, P)dy = \mathcal{N}\left(x; Fm; Q + FPF^T\right)$$

# The Kalman filter

Identity 2: Given $H, R, m, P$ of suitable dimensions, and $R$ and $P$ are positive definite, then

$$\mathcal{N}(z; Hx, R)\mathcal{N}(x; m, P) = q(z)\mathcal{N}(x; \bar{m}, \bar{P})$$

where

$$\begin{aligned} q(z) &= \mathcal{N}\left(z; HM, R + HPH^T\right) \\ \bar{m} &= m + K(z - Hm) \\ \bar{P} &= (I - KH)P \\ K &= PH^T\left(HPH^T + R\right)^{-1}. \end{aligned}$$

# The Kalman filter

The time-update equation for the Kalman filter

$$
\begin{aligned}
p_{k|k-1}\left(x_k \mid z_{1:k-1}\right) &= \int f_{k|k-1}\left(x_k \mid x\right) p_{k-1}\left(x \mid z_{1:k-1}\right) dx, \\
&= \int \mathcal{N}\left(x_k; F_{k-1}x, Q_{k-1}\right) \mathcal{N}\left(x; \hat{x}_{k-1}, P_{k-1}\right) dx \\
&= \mathcal{N}\left(x_k; \hat{x}_{k|k-1}, P_{k|k-1}\right),
\end{aligned}
$$

# The Kalman filter

The predicted mean and covariance are calculated with

$$\hat{x}_{k|k-1} = F_{k-1}\hat{x}_{k-1|k-1}$$
$$P_{k|k-1} = Q_{k-1} + F_{k-1}P_{k-1|k-1}F_{k-1}^T$$

# The Kalman filter

The measurement update for the Kalman filter

$$p_k\left(x_k \mid z_{1:k}\right) = \frac{g_k\left(z_k \mid x_k\right) p_{k|k-1}\left(x_k \mid z_{1:k-1}\right)}{\int g_k\left(z_k \mid x\right) p_{k|k-1}\left(x \mid z_{1:k-1}\right) dx}.$$

$$x_{\hat{k}|k} = \hat{x}_{k|k-1} + K_k\left(z_k - H_k \hat{x}_{k|k-1}\right)$$

$$P_{k|k} = \left(I - K_k H_k\right) P_{k|k-1},$$

where

$$K_k = P_{k|k-1} H_k^T \left(H_k P_{k|k-1} H_k^T + R_k\right)^{-1}$$

is known as the Kalman gain.

# Extended Kalman filter

In the extended Kalman filter, the state transition and observation models don't need to be linear functions of the state but may instead be differentiable functions.

$$\boldsymbol{x}_k = f(\boldsymbol{x}_{k-1}) + \boldsymbol{w}_{k-1}$$
$$\boldsymbol{z}_k = h(\boldsymbol{x}_k) + \boldsymbol{v}_k$$

Here $\mathbf{w}_k$ and $\mathbf{v}_k$ are the process and observation noises which are both assumed to be zero mean multivariate Gaussian noises with covariance $\mathbf{Q}_k$ and $\mathbf{R}_k$ respectively.

# Extended Kalman filter

The function $f$ is used to compute the predicted state from the previous estimate. However, $f$ cannot be applied to the covariance directly. Instead a matrix of partial derivatives (the Jacobian) is computed.

# Extended Kalman filter

The Jacobian is evaluated with current predicted states. The Jacobian matrix is used in the Kalman filter equations. This process linearizes the non-linear function around the current estimate.

# Extended Kalman filter

The state transition and observation matrices are defined to be the following Jacobians

$$\boldsymbol{F}_k = \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\hat{\boldsymbol{x}}_{k-1|k-1}}$$

$$\boldsymbol{H}_k = \left.\frac{\partial h}{\partial \boldsymbol{x}}\right|_{\hat{\boldsymbol{x}}_{k|k-1}}$$

# Extended Kalman filter

$$
\begin{aligned}
\hat{\boldsymbol{x}}_{k|k-1} &= f\left(\hat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{u}_{k-1}\right) \\
\boldsymbol{P}_{k|k-1} &= \boldsymbol{F}_k \boldsymbol{P}_{k-1|k-1} \boldsymbol{F}_k^T + \boldsymbol{Q}_{k-1}
\end{aligned}
$$

# Extended Kalman filter

$$\tilde{\boldsymbol{y}}_k = \boldsymbol{z}_k - h\left(\hat{\boldsymbol{x}}_{k|k-1}\right)$$
$$\boldsymbol{S}_k = \boldsymbol{H}_k \boldsymbol{P}_{k|k-1} \boldsymbol{H}_k^T + \boldsymbol{R}_k$$
$$\boldsymbol{K}_k = \boldsymbol{P}_{k|k-1} \boldsymbol{H}_k^T \boldsymbol{S}_k^{-1}$$
$$\hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k \tilde{\boldsymbol{y}}_k$$
$$\boldsymbol{P}_{k|k} = \left(\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{H}_k\right) \boldsymbol{P}_{k|k-1}$$

# Extended Kalman filter

where the state transition and observation matrices are defined to be the following Jacobians

$$\boldsymbol{F}_k = \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\hat{\boldsymbol{x}}_{k-1|k-1}}$$

$$\boldsymbol{H}_k = \left.\frac{\partial h}{\partial \boldsymbol{x}}\right|_{\hat{\boldsymbol{x}}_{k|k-1}}$$

# Extended Kalman filter

$$\boldsymbol{z} = \begin{bmatrix} r \\ \theta \end{bmatrix}$$
$$= \begin{bmatrix} \sqrt{x^2 + y^2} \\ \tan^{-1}(y/x) \end{bmatrix}$$
$$H = \frac{\partial z}{\partial \boldsymbol{x}} = ?$$

# Extended Kalman filter

What is the Jacobian?

$$H = \begin{bmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial \dot{x}} & \frac{\partial r}{\partial y} & \frac{\partial r}{\partial \dot{y}} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial \dot{x}} & \frac{\partial \theta}{\partial y} & \frac{\partial \theta}{\partial \dot{y}} \end{bmatrix};$$

$$\begin{aligned} z &= \begin{bmatrix} r \\ \theta \end{bmatrix} \\ &= \begin{bmatrix} \sqrt{x^2 + y^2} \\ \tan^{-1}(y/x) \end{bmatrix} \end{aligned}$$

$$H = \frac{\partial \boldsymbol{z}}{\partial \boldsymbol{x}} = ?$$

# Extended Kalman filter

What is the Jacobian?
Jacobian used in the Taylor series expansion looks like ...

$$H = \begin{bmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial \dot{x}} & \frac{\partial r}{\partial y} & \frac{\partial r}{\partial \dot{y}} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial \dot{x}} & \frac{\partial \theta}{\partial y} & \frac{\partial \theta}{\partial \dot{y}} \end{bmatrix} = \begin{bmatrix} \cos(\theta) & 0 & \sin(\theta) & 0 \\ -\sin(\theta)/r & 0 & \cos(\theta)/r & 0 \end{bmatrix}$$

# Particle filter

Recall the Bayesian filter:

$$1)\ p_{k|k-1}\left(x_k \mid z_1,\ldots,z_{k-1}\right) =$$
$$\int f_{k|k-1}\left(x_k \mid x_{k-1}\right) p_{k-1}\left(x_{k-1} \mid z_1,\ldots,z_{k-1}\right) dx_{k-1}$$
$$2)\ p_k\left(x_k \mid z_1,\ldots,z_{k-1}\right) \propto g_k\left(z_k \mid x_k\right) p_{k|k-1}\left(x_k \mid z_1,\ldots,z_{k-1}\right).$$

When the posterior density $p_{k-1}$, Markov transition function, and likelihood are non-Gaussian or the state and observation models are nonlinear, then the condition for the Kalman filter are no longer valid.

# Particle filter

One commonly used technique is the "particle filter", which represents (approximates) the posterior density by particles. We have $N$ independent samples for $p_{k-1}$ at time step $k-1$, i.e. $\left\{x_{k-1}^{(1)}, \cdots, x_{k=1}^{(N)}\right\}$. An empirical estimate of $p_{k-1}$ is given by

$$p_{k-1}^{N}\left(x \mid z_1, \ldots x_{k-1}\right) \approx \frac{1}{N}\sum_{i=1}^{N} \delta_{x_{k-1}^{(i)}}\left(x\right)$$

where

$$\delta_{x_{k-1}^{(i)}}\left(x_{k-1}\right) = \begin{cases} \infty & \text{if } x_{k-1}^{(i)} = x_{k-1} \\ 0 & \text{otherwise.} \end{cases}$$

# Particle filter

We can approximate integrands of the form

$$I(f) = \int f(x)\, p_{k-1}(x)\, dx$$

with

$$I_N(f) = \int f(x) \sum_{i=1}^{n} \delta_{x_{k-1}^{(i)}}(x)\, dx$$

which is equal to

$$\frac{1}{N} \sum_{i=1}^{N} f\left(x_{k-1}^{i}\right)$$

# Particle filter

**Importance Sampling:**
Let us introduce an arbitrary "proposal" distribution, $\pi_{k-1}(x_{k-1}|z_{1:k-1})$. Then, if we can sample $N$ particles $x_{k-1}^{(i)}, i = 1, \ldots, N$ from $\pi_{k-1}$, and set

$$w(x) = \frac{p_{k-1}(x|z_1, \ldots z_{k-1})}{\pi_{k-1}(x|z_1, \ldots z_{k-1})}. \tag{1}$$

Then we find

$$I(f) = \frac{\int f(x)w(x)\pi_{k-1}(x_1 \mid z_1, \ldots z_{k-1})\, dx}{\int w(x)\pi_{k-1}(x \mid z_1, \ldots z_{k-1})\, dx}$$

# The Bootstrap filter

Hence

$$I(f) = \left[\frac{1}{N}\sum_{i=1}^{N} f\left(x_{i-1}^{(i)}\right) w\left(x_{k-1}^{(i)}\right)\right] \Big/ \left[\frac{1}{N}\sum_{i=1}^{N} w\left(x_{k-1}^{i}\right)\right]$$

$$= \sum_{i=1}^{N} f\left(x_{k-1}^{(i)}\right) \tilde{w}_{k-1}\left(x_{k-1}^{(i)}\right),$$

where the weights $\tilde{w}_{k-1}^{(i)}$ have been normalised.

# The Bootstrap filter

There are different ways of creating particle filters. The "Bootstrap filter" is the simplest and most used particle filter.

- 1) **Initialize** at time step 0 For $i = 1 \ldots N_1$ sample $x_0^{(i)} \sim p_0$. Then set time to $k = 1$.
- 2) **Importance sampling:** For $i = 1, \ldots N$ sample $x_k^{(i)} \sim f_{k|k-1}\left(\cdot \mid x_{k-1}^{(i)}\right)$.
- 3) **Update:** Set the weights to be $w_k^{(i)} = g_k\left(z_k \mid x_k^{(i)}\right)$.
- 4) **Resample** $N$ particles $x_k^{(i)}, i = 1, \ldots, N$ according to the weights.
- 5) Set $k = k + 1$ and go to step 2.