AUTOMATIC CONTROL
KTH
Applied Estimation EL2320
Exam 9:00-13:00 January 14, 2015

Aids: None, no books, no notes, nor calculators

Observe:

- Name and person number on every page
- Answers should be in English (or Swedish)
- Only write on one side of the sheets
- Specify the total number of handed in pages on the cover
- Be careful to label each answer with the question number and letter
- Be careful to only state things you know are true as one incorrect statement can negate an otherwise correct or partially correct answer.
- That said also be sure to answer the question asked.
- Motivate answer and clearly state any additional assumptions you may need to make.

Results: The results will be available at STEX, in a few weeks
Studerandexpeditionen, Osquldasv. 10. **Responsible:** John Folkesson, 08-790-6201

Good to know (what all the symbols mean you should be able to figure, this is just in case you forgot how to put them together):

$$G(\mathbf{x}, \mu, \Sigma) = \frac{1}{\sqrt{|2\pi\Sigma|}} \exp \frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) \tag{1}$$

Typical Non-linear system equations:

$\mathbf{x}_t = \mathbf{g}(\mathbf{u}_t, \mathbf{x}_{t-1}) + \varepsilon_t$

$\mathbf{z}_t = \mathbf{h}(\mathbf{x}_t) + \delta_t$

'Kalman Gain'

$K_t = \bar{\Sigma}_t H_t^T (H_t \bar{\Sigma}_t H_t^T + Q_t)^{-1}$

Predict phase:

$\bar{\Sigma}_t = R_t + G_t \Sigma_{t-1} G_t^T$

$\bar{\mu}_{\mathbf{t}} = \mathbf{g}(\mathbf{u}_t, \mu_{t-1})$

Update:

$\Sigma_t = \bar{\Sigma}_t - K_t H_t \bar{\Sigma}_t$

$\mu_t = \bar{\mu}_{\mathbf{t}} + K_t \eta_t$

Questions

1. (5p) For a joint probability given by:

$P(A,B)$	$B=0$	$B=1$	$B=2$	$B=3$
$A=0$	0.10	?	0.15	0.00
$A=1$	0.05	0.10	0.00	0.10
$A=2$	0.00	0.05	0.05	0.05
$A=3$	0.10	0.10	0.10	0.00

a) What is the value of the '?' in the table? (1p) ans. 0.05

b) What is the conditional density function of B given A, $P(B|A)$? Give the answer as a table. You can use fractions, such as for example $\frac{0.10}{0.70}$ (1p) ans.

$P(B\|A)$	$B=0$	$B=1$	$B=2$	$B=3$
$A=0$	$\frac{0.10}{0.30}$	$\frac{0.05}{0.30}$	$\frac{0.15}{0.30}$	0.00
$A=1$	$\frac{0.05}{0.25}$	$\frac{0.10}{0.25}$	0.00	$\frac{0.10}{0.25}$
$A=2$	0.00	$\frac{0.05}{0.15}$	$\frac{0.05}{0.15}$	$\frac{0.05}{0.15}$
$A=3$	$\frac{0.10}{0.30}$	$\frac{0.10}{0.30}$	$\frac{0.10}{0.30}$	0.00

c) What is the probability $P(B)$, for each value of B? (1p)

$P(A,B)$	$B=0$	$B=1$	$B=2$	$B=3$
	0.25	0.30	0.30	0.15

d) What is the expected value of B? (1p)
ans. $.3 + .6 + .45 = 1.35$

e) What is the variance of B given that A=0? (1p)
ans. $(0.05 + 0.6)/0.3 - [(0.5 + 0.3)/0.3]^2 = 65/30 - [7/6]^2$
$= 13/6 - 49/36 = (78 - 49)/36 = 29/36$

2. (4p) As is well known, all roads lead to Rome. You find yourself on a road but have no idea which direction along it is towards Rome. You flip a coin and start driving in a direction that you had assigned to the flip outcome. Let $R \in \{0, 1\}$ be a random variable that indicates your direction with $R = 1$ being towards Rome and $R = 0$ being away from Rome.

   a) Assuming all roads have one end towards Rome and one away from it, what is the probability that you are heading towards Rome, $P(R = 1)$? (1p)

   ans. 0.50

   b) You notice that the traffic density is higher in the direction you are driving. You know that at this time of day $\frac{1}{2}$ of the days, on average, the traffic is greater going towards Rome while $\frac{1}{4}$ the days it is greater going the other way and the rest of the days there is no difference in traffic with direction. What is your belief that $R = 1$ now? Include a few lines of probablistic reasoning starting from the given information translated to mathematical notations. (3p)

   ans. Let T be the measurement of the traffic such that T=1 indicates heavier traffic in the direction of travel, T=-1 indicate that is in the other direction, and T=0 indicate no difference. Then the above information translates to:

   $P(T = 1|R = 1) = 0.5$ and $P(T = 1|R = 0) = 0.25$

   $P(R = 1|T = 1) = P(T = 1|R = 1)P(R = 1)/P(T = 1)$
   $= (0.5)(0.5)/[(0.5)(0.5) + (0.25)(0.5)]$
   $= 0.5/0.75 = 0.67$

3. (13p) This quesion is on the Kalman Filter. (Not the EKF)

   a) Kalman filters make estimates of dynamic states. What are the limitations/assumptions on the dynamics (aka process model)? (2p)

   ans. The process model is linear the noise is Gaussian and the process is first order Markov, that is all the future states and process noise at time t are independent of the previous data given the state at time t.

b) Kalman filters use measurements to improve the estimates. What are the limitations/assumptions about the measurements? (2p)
ans. The measurement model is linear the noise is Gaussian and the process is first order Markov, that is all the measurements and noise at time t and after are independent of the previous data given the state at time t.

c) If all the assumptions in (a) and and (b) are true what is shape of the posteriori distribution? (2p)
ans. Gaussian

d) All the assumptions in (a) and and (b) are true and we get a measurement that indicates that the error in our estimated mean is unreasonably large. We use that measurement to improve our estimated mean according to the KF equations. Would we have been able to, using some other estimation method, get a better estimate if we had not marginalized out all the earlier states? Explain why/why not. (2p) ans. No since the assumptions imply that the true belief is Gaussian and exactly equal to the Kalman estimate. There would have been no approximations done and all information was used, so there was no chance of improving the accuracy. We might happen to sometimes get numbers closer to the true state by chance but the KF is the optimal estimator.

e) There are two phases of a Kalman Filter, predict and update, which of these phases apply Bayesian inference? (2p)
ans. Update

f) If we compensate for measurement modeling errors by increasing the measurement noise in what sense might our series of estimates improve? In what sense might it become worse? (3p)
ans.
Our estimates might be more consistent as the model error would otherwise lead to overconfidence. Over confidence could even lead to convergence to the wrong state. However, our esimates would converge more slowly as a result.

4. (5p) The example of robot localization was used in the labs, lectures and book.

   a) Give an example (simple is good) of what the state vector might consist of for robot localization. (1p)

   ans.
   The state might be (x, y, θ) where x and y are the Cartesian coordinates of the robot and θ is the angle of its heading relative to the x axis.

   b) Based on your example specify a possible nonlinear dynamic model including Gaussian noise and explaining where all parameters would come from. (2p)

   ans.
   $x_{t+1} = x_t + \Delta s \cos(\theta_t) + \eta_x$
   $y_{t+1} = y_t + \Delta s \sin(\theta_t) + \eta_y$
   $\theta_{t+1} = \theta_t + \Delta\theta + \eta_\theta$

   The η are Gaussian noise with zero mean and a covariance estimated based on for example measurements of the actual system. The Δs and $\Delta\theta$ can be gotten from measurements of the wheel rotations.

   c) What is the 'G' Jacobian matrix needed for a EKF prediction based on your model in (b). (1p)

   ans.

$$G_{t+1} = \begin{bmatrix} 1 & 0 & -\Delta s \sin\theta_t \\ 0 & 1 & \Delta s \cos\theta_t \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

   d) Describe any problems that might occur in general with estimating the state over a longer time using a Kalman Filter based solely on a dynamic model without any other measurements. (1p)

   ans. The estimation error will often increase without bound unless there are measurements to limit it.

5. (13 p) These questions are on the particle filter, PF. The PF algorithm recursively estimates the belief of the state of a dynamic system consisting of a process model and a measurement model.

a) What are the inputs (and the outputs) of the recursive algorithm. That is how is the belief represented at time t. (1p)

ans. A set of M 'particles', $\{\mathbf{x}_t^i\}$. Each particle is a point in the state space. It is hoped that these points are drawn from the priori/posteori distributions.

b) Outline the steps of the recursive algorthim from input to output. (3p)

ans.

- 1. for each particle $\mathbf{x}_t^i$ draw a new sample:
  $\mathbf{x}_{t+1}^i \sim p(\mathbf{x}_{t+1}^i | \mathbf{x}_t^i, \mathbf{u}_t)$
- 2. compute a weight for each particle $w^i = p(z_{t+1}|x_{t+1}^i)$.
- 3. normalize the sum of weights to 1.0.
- 4. resample the particles with replacement by selecting with probablity given by the weights.

c) What assumptions are made on the form of the *a priori* and *a posteori* distributions? (1p)

ans.

Saying that there are no assumptions is accepted but there are better answers: The markov property. Also the support of these must equal the support of the proposal distribution. We also need to be able to compute the weights.

d) What assumptions are made on the process noise and measurement noise? (2p)
ans.
Markov property. They should be independent given the latest state.

e) Describe particle deprivation. What is it and how does it happen? (2 p)
ans.
It is when there are not sufficient particles in some region of the state space with significant probability. It results from the random sampling of a finite number of particles.

f) Why do some variations on the PF use different proposal distributions? What properties make one proposal distribution better than another? (2 p)

ans. They typically are hoping to reduce particle deprivation by focusing the particle set to more likely regions before importance resampling. A good proposal distribution matches the posteriori approximately and gives weights that can be computed relatively easily.

g) What is the advantage of increasing the modeled noise during the diffusion step? (2p)

ans. This will cause greater spreading of the particles and give greater exploration of the possible states. This can then slow particle deprivation.

6. (10p)

a) How do the Kalman Filter and Extended Kalman Filter differ in the types of dynamic models? (2p)

ans. KF must have a linear model while EKF can be non-linear.

b) How do the Unscented Kalman Filter and Extended Kalman Filter differ in the way they approximate nonlinearities in the process model? (2p)

ans. The UKF passes a set of carefully chosen points through the non-linearity while the EKF passes only one but approximates the function linearly to obtain the distributions 'spread'.

c) How do the 'vanilla' and Rao-Blackwellized Particle Filter differ in how they represent the 'belief'? (2p)

ans. The vanilla PF uses a set of particles drawn from the postiori over the full state space while the RB PF uses particles drawn from the the postiori over a subspace of the state space and for each of those a paremetric distribution over the rest of the state space.

d) What is the difference in the update steps betwen the Iterative Extended Kalman Filter and Extended Kalman Filter? (2p) ans. The iEKF does the EKF update repeatedly each time linearizing around the previous step's solution. The EKF only does one linearization and update based on linearizing around the predicted state.

e) Explain what is meant by the Markov assumption needed for the EKF and PF? (2p)

ans. All data after time t is independent of the data before time t given the state at time t.