Thesis Filter Basics

Arne Siebenmorgen

October 2017

1 Opper

Filter Equation:

$$P(z,\theta|x_{t+1}) = \alpha \cdot P(x_{t+1}|\theta) \cdot \sum_{z'} P(z|z') \cdot P(z',\theta|x_{1:t}) \tag{1}$$

Approximation $P(z,\theta) = P(z)P(\theta)$
Einsetzten und Integration über θ:

$$P(z|x_{t+1}) = \alpha \cdot \sum_{z'} P(z|z')P_t(z') \cdot \int P(x_{t+1}|\theta)P(\theta|x_{1:t})d\theta \tag{2}$$

$$= \alpha \cdot \sum_{z'} P(z|z')P_t(z') \cdot \int P(x_{t+1}|\theta_z)P(\theta_z|x_{1:t})d\theta_z \tag{3}$$

Approximation $P(\theta) = P(\theta_1)P(\theta_2)$
Einsetzen und über z summieren:

$$P(\theta|x_{t+1}) = \alpha \cdot P(\theta|x_{1:t}) \cdot \sum_{z} P(x_{t+1}|\theta) \sum_{z'} P(z|z')P_t(z') \tag{4}$$

$$= \alpha \cdot P(\theta|x_{1:t}) \cdot \left(\sum_{z'} P(z_1|z')P_t(z')P(x_{t+1}|\theta) + \sum_{z'} P(z_2|z')P_t(z') \int P(\theta_2|x_{1:t})P(x_{t+1}|\theta_2)\, d\theta_2 \right) \tag{5}$$

Berechnung vom Integral $\int P(x_{t+1}|\theta_z)P(\theta_z|x_{1:t})d\theta_z$

$$\int P(x_{t+1}|\theta_z)P(\theta_z|x_{1:t})d\theta_z = \int \mathcal{N}(x_{t+1}|\mu_\theta, 1)\mathcal{N}(\mu_\theta|\mu_{1:t}, 1)d\mu_\theta \tag{6}$$

$$= \int \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}(x_{t+1} - \mu_\theta)^2) \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}(\mu_\theta - \mu_{1:t})^2)d\mu_\theta \tag{7}$$

$$= \int \frac{1}{2\pi} \exp(-\frac{1}{2}(x_{t+1}^2 - 2x_{t+1}\mu_\theta + 2\mu_\theta^2 - 2\mu_\theta\mu_{1:t} + \mu_{1:t}^2))d\mu_\theta \tag{8}$$

$$= \int \frac{1}{2\pi} \exp(-\frac{1}{2}(2(\mu_\theta^2 - \mu_\theta^2(x_{t+1} + \mu_{1:t}) + \frac{x_{t+1}^2}{2} + \frac{\mu_{1:t}^2}{2})))d\mu_\theta \tag{9}$$

$$= \int \frac{1}{2\pi} \exp(-\frac{1}{4}(\mu_\theta^2 - 2\mu_\theta\frac{x_{t+1} + \mu_{1:t}^2}{2} + \frac{x_{t+1}}{2} + \frac{\mu_{1:t}}{2} + \frac{x_{t+1} + \mu_{1:t}^2}{4} - \frac{x_{t+1} + \mu_{1:t}^2}{4}))d\mu_\theta \tag{10}$$

$$= \int \frac{1}{2\pi} \exp(-\frac{1}{4}(\mu_\theta^2 - 2\mu_\theta\frac{x_{t+1} + \mu_{1:t}^2}{2} + \frac{x_{t+1}}{2} + \frac{\mu_{1:t}}{2} + \frac{x_{t+1} + \mu_{1:t}^2}{4} - \frac{x_{t+1} + \mu_{1:t}^2}{4}))d\mu_\theta \tag{11}$$

$$= \int \frac{1}{2\pi} \exp(-\frac{1}{2}\frac{(\mu_\theta - \frac{(x+\mu_{1:t})^2}{2})^2}{1/2}) \exp(\frac{(x + \mu_{1:t})^2}{4} - \frac{x_{t+1}^2}{2} - \frac{\mu_{1:t}^2}{2})d\mu_\theta \tag{12}$$

$$= \frac{\sqrt{\pi}}{2\pi} \exp(\frac{(x + \mu_{1:t})^2}{4} - \frac{x_{t+1}^2}{2} - \frac{\mu_{1:t}^2}{2}) \int \frac{1}{2\pi\frac{1}{2}} \exp(-\frac{1}{2}\frac{(\mu_\theta - \frac{(x+\mu_{1:t})^2}{2})^2}{1/2})d\mu_\theta \tag{13}$$

$$= \frac{\sqrt{\pi}}{2\pi} \exp(\frac{(x + \mu_{1:t})^2}{4} - \frac{x_{t+1}^2}{2} - \frac{\mu_{1:t}^2}{2}) \tag{14}$$

2 Mit Sebastian

Filter Equation

$$P(z_{t+1}, \theta_{t+1}|x_{t+1}) = \alpha \cdot P(x_{t+1}|\theta) \cdot \sum_{z_t} P(z_{t+1}|z_t) \cdot P(z_t, \theta|x_{1:t}) \tag{15}$$

θ is now in z:

$$P(z_{t+1}|x_{1:t+1}) = \alpha \cdot P(x_{t+1}|z_{t+1}) \cdot \sum_{z_t} P(z_{t+1}|z_t) \cdot P(z_t|x_{1:t}) \tag{16}$$

$$P(z_{t+1}|x_{1:t+1}) = \alpha \cdot P(x_{t+1}|z_{t+1}) \cdot \sum_{z_t} P(z_{t+1}, z_t|x_{1:t}) \tag{17}$$

$$P(z_{t+1}|x_{1:t+1}) = \alpha \cdot P(x_{t+1}|z_{t+1}) \cdot P(z_{t+1}|x_{1:t}) \tag{18}$$

$$P(z_{t+1}|x_{1:t+1}) = \alpha \cdot P(x_{t+1}, z_{t+1}|x_{1:t}) \tag{19}$$

$$P(z_{t+1}|x_{t+1}, x_{1:t}) = \alpha \cdot P(x_{t+1}, z_{t+1}|x_{1:t}) \tag{20}$$

$$\frac{P(z_{t+1}, x_{t+1}|x_{1:t})}{P(x_{t+1}|x_{1:t})} = P(x_{t+1}, z_{t+1}|x_{1:t}) \tag{21}$$

$$P(z_{t+1}|x_{1:t+1}) = \frac{P(x_{t+1}|z_{t+1}) \cdot P(z_{t+1}|x_{1:t})}{P(x_{t+1}|x_{1:t})} \tag{22}$$

3 Andreas

Target:

$$P(z_{t+1}, \theta_1, \theta_2|x_{1:t+1}) = \frac{P(x_{t+1}, z_{t+1}, \theta_1, \theta_2|x_{1:t})}{P(x_{t+1}|x_{1:t})} \tag{23}$$

$$P(z_{t+1}, \theta_1, \theta_2|x_{1:t+1}) = \alpha \cdot P(x_{t+1}, z_{t+1}, \theta_1, \theta_2|x_{1:t}) \tag{24}$$

$$P(z_{t+1}, \theta_1, \theta_2|x_{1:t+1}) = \alpha \cdot P(x_{t+1}|z_{t+1}, \theta_1, \theta_2) \cdot P(z_{t+1}, \theta_1, \theta_2|x_{1:t}) \tag{25}$$

$$\begin{split} P(z_{t+1}, \theta_1, \theta_2|x_{1:t+1}) = {} & \alpha \cdot P(x_{t+1}|z_{t+1}, \theta_1, \theta_2) \cdot \\ & \sum_{z_t} P(z_{t+1}|z_t) \cdot P(z_t, \theta_1, \theta_2|x_{1:t}) \end{split} \tag{26}$$

$$\begin{split} P(z_{t+1}, \theta_1, \theta_2|x_{1:t+1}) = {} & \alpha \cdot P(x_{t+1}|z_{t+1}, \theta_1, \theta_2) \cdot \\ & \sum_{z_t} P(z_{t+1}|z_t) \cdot P(z_t|x_{1:t}) \cdot P(\theta_1|x_{1:t}) \cdot P(\theta_2|x_{1:t}) \end{split} \tag{27}$$

For each Iteration:

$$P(z_{t+1}, \theta_1, \theta_2|x_{1:t+1}) = \alpha \cdot P(x_{t+1}, z_{t+1}, \theta_1, \theta_2|x_{1:t}) \tag{28}$$

$$z_{t+1} = 1 : P(\theta_1|x_{1:t+1}) = \alpha \cdot P(x_{t+1}|\theta_1) \cdot P(\theta_1|x_{1:t}) \tag{29}$$

$$z_{t+1} = 2 : P(\theta_2|x_{1:t+1}) = \alpha \cdot P(x_{t+1}|\theta_2) \cdot P(\theta_2|x_{1:t}) \tag{30}$$

To update θ:

$$\mu_{t+1} = \frac{1}{\sigma_{t+1}^2} \left(\frac{\sum_t x}{\sigma_{t+1}^2} + \frac{\mu_t}{\sigma_t^2} \right) \tag{31}$$

$$\frac{1}{\sigma_{t+1}^2} = \frac{1}{\sigma_t^2} + \frac{t}{\sigma_t^2} \tag{32}$$

4 Conjugate Gaussian

4.1 Multiplication of two Gaussian

For the update at each iteration $P(\theta|x_{1:t+1}) = \alpha \cdot P(x_{t+1}|\theta) \cdot P(\theta|x_{1:t})$ needs to be computed. $P(x_{t+1}|\theta)$ and $P(\theta|x_{1:t})$ are both assumed to be Gaussian. This leads to the following derivation of updates:
Let f and g be Normal PDF

$$f(x) = N(x|\mu_1, \sigma_1), \quad g(x) = N(x|\mu_2, \sigma_2) \tag{33}$$

$$f(x) \cdot g(x) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(- \left(\frac{(x-\mu_1)^2}{2\sigma_1^2} + \frac{(x-\mu_2)^2}{2\sigma_2^2} \right) \right) \tag{34}$$

Looking into the exponent

$$\frac{(x-\mu_1)^2}{2\sigma_1^2} + \frac{(x-\mu_2)^2}{2\sigma_2^2} = \frac{(\sigma_1^2+\sigma_2^2)x^2 - 2(\mu_1\sigma_2^2+\mu_2\sigma_1^2)x + \mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2}{2\sigma_2^2\sigma_1^2} \tag{35}$$

$$= \frac{x^2 - 2\frac{\mu_1\sigma_2^2+\mu_2\sigma_1^2}{\sigma_1^2+\sigma_2^2}x + \frac{\mu_1^2\sigma_2^2+\mu_2^2\sigma_1^2}{\sigma_1^2+\sigma_2^2}}{2\frac{\sigma_2^2\sigma_1^2}{\sigma_1^2+\sigma_2^2}} \tag{36}$$

The exponent of a regular Gaussian can be rewritten as the following:

$$\frac{(\mu-x)^2}{2\sigma^2} \tag{37}$$

$$= \frac{\mu^2 - \mu x + x^2}{2\sigma^2} \tag{38}$$

With that we can infer the new mean and variance of the scaled Gaussian $P(\theta|x_{1:t+1})$

$$\mu = \frac{\mu_1\sigma_2^2 + \mu_2\sigma_1^2}{\sigma_1^2+\sigma_2^2} \tag{39}$$

$$\sigma^2 = \frac{\sigma_2^2\sigma_1^2}{\sigma_1^2+\sigma_2^2} \tag{40}$$

4.2 Derivation from Kevin P. Murphy

The problem with this derivation is, that it leads to the correct updates for θ as defined by Andreas, but the approach is not equivalent to our.

4.2.1 Likelihood

D is the data $x_1, ..., x_t$

$$
\begin{aligned}
P(D|\mu, \sigma^2) &= P(x_{1:t}|\mu, \sigma^2) && (41)\\
&= \prod^{t} P(x_i|\mu, \sigma^2) && (42)\\
&= (2\pi\sigma^2)^{-\frac{n}{2}} \cdot \exp(-\frac{1}{2\sigma^2} \sum (x_i - \mu)^2) && (43)
\end{aligned}
$$

4.2.2 Prior

$$
P(\mu|\mu_0, \sigma^2) \tag{44}
$$

4.3 Posterior

$$
\begin{aligned}
P(\mu|x_{1:t}) &= \alpha \cdot P(D|\mu, \sigma^2) P(\mu|\mu_0, \sigma^2) && (45)\\
&= \alpha \cdot \exp(-\frac{1}{2\sigma^2} \sum (x_i - \mu)^2) \exp(-\frac{1}{2\sigma^2}(\mu - \mu_0)^2) && (46)\\
&= \alpha \cdot \exp(-\frac{1}{2\sigma^2} \sum (x_i - \mu)^2) \exp(-\frac{1}{2\sigma^2}(\mu - \mu_0)^2) && (47)\\
&= \alpha \cdot \exp(-\frac{\mu^2}{2}(\frac{1}{\sigma^2} + \frac{n}{\sigma^2}) + \mu(\frac{\mu_0}{\sigma_0^2} + \frac{\sum x_i}{\sigma^2}) - (\frac{\mu_0^2}{2\sigma_0^2} + \frac{\sum x_i^2}{2\sigma^2})) && (48)
\end{aligned}
$$

Now we have to match the coefficients to a regular Gaussian:

$$
\exp(\frac{1}{2\sigma_n^2}(\mu - \mu_2)^2) = \exp(\frac{1}{2\sigma_n^2}(\mu^2 - 2\mu\mu_n + 2\mu^2)) \tag{49}
$$

Matching this to equation [?] yields:

$$
\begin{aligned}
-\frac{\mu^2}{2\sigma_n^2} &= -\frac{\mu^2}{2}(\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}) && (50)\\
-\frac{1}{\sigma_n^2} &= \frac{1}{\sigma_0^2} + \frac{n}{\sigma^2} && (51)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{2\mu\mu_n}{2\sigma_n^2} &= \mu(\frac{\mu_0}{\sigma_0^2} + \frac{\sum x_i}{\sigma^2}) && (52)\\
\frac{\mu_n}{\sigma_n^2} &= \frac{\mu_0}{\sigma_0^2} + \frac{\sum x_i}{\sigma^2} && (53)
\end{aligned}
$$

Algorithm 1 Digital Assignment Two State Gaussian

while new data **do**
  $x_{t+1} \leftarrow newdata$
  $likelihoodState_0 \leftarrow normal(x_{t+1}|\theta_0, 1)$
  $likelihoodState_1 \leftarrow normal(x_{t+1}|\theta_1, 1)$
  $\tilde{z}_0 \leftarrow likelihoodState_0 \cdot (P(z_0|z_0)P(z_0|x_{1:t}) + P(z_0|z_1)P(z_1|x_{1:t}))$
  $\tilde{z}_1 \leftarrow likelihoodState_1 \cdot (P(z_1|z_0)P(z_0|x_{1:t}) + P(z_1|z_1)P(z_1|x_{1:t}))$
  $z_0 \leftarrow z_0/(z_0 + z_1)$
  $z_1 \leftarrow z_1/(z_0 + z_1)$
  if $z_0 \geq z_1$ **then**
    $\sigma_0 \leftarrow \frac{1}{1/\theta_0+t}$
    $\theta_0 \leftarrow \frac{\theta_0/\sigma_0+\sum x/\sigma_0}{1/\sigma_0+t/\sigma_0}$
  else
    $\sigma_1 \leftarrow \frac{1}{1/\theta_1+t}$
    $\theta_1 \leftarrow \frac{\theta_1/\sigma_1+\sum x/\sigma_1}{1/\sigma_1+t/\sigma_1}$

5 Algorithms

5.1 Evaluation Methods

1. How long does it take to find the true mean
2. False Positive rate of predicting z
3. How does the algorithm reacts to far off prior values to start with
4. How does the algorithm reacts to the wrong number of states

6 Estimating Markov Transition Matrix

Since the time series is discrete valued, one can estimate the transition probabilities by the sample proportions. Let Y_t be the state of the process at time t, M be the transition matrix then

$$M_{ij} = P(Y_t = j|Y_{t-1} = i) \tag{54}$$

Since this is a markov chain, this probability depends only on Y_{t1}, so it can be estimated by the sample proportion. Let n_{ik} be the number of times that the process moved from state i to k. Then,

$$\hat{M_{ij}} = \frac{n_{ij}}{\sum_{k=1}^{m} n_{ik}} \tag{55}$$

where m is the number of possible states. The denominator, $\sum_{k=1}^{m} n_{ik}$, is the total number of movements out of state i. Estimating the entries in this way actually corresponds to the maximum likelihood estimator of the transition matrix, viewing the outcomes as multinomial, conditioned on Y_{t-1}.