University of Tehran

June 2021

Extended Kalman Filter using Normalizing Flow

Ali Izadi

۱- مروری بر مقاله Normalizing Kalman Filters for Multivariate Time Series Analysis

۲- تلاش برای اضافه کردن Normalizing Flow به Extended Kalman Filters

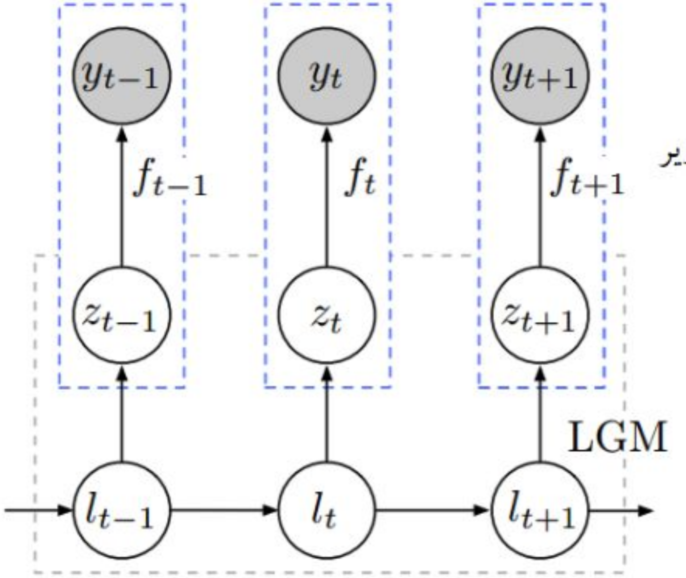
مدل kalman filter

روابط kalman filter برای زمانی که تابع بین stateها و observationها غیر خطی است در زیر آمده است:

$$l_1 \sim \mathcal{N}(\mu_1,\, \sigma_1^2)\,.$$

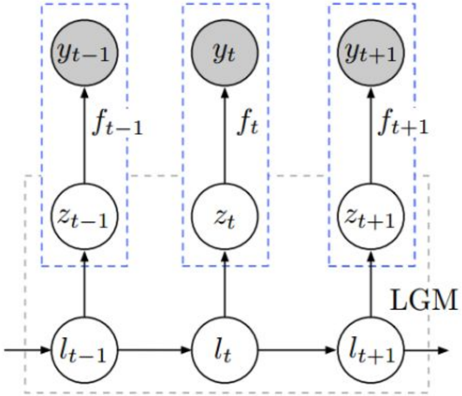
$$l_t = F_t l_{t-1} + \epsilon_t \qquad \epsilon_t \sim \mathcal{N}(0,\, \sigma_{1t}^2)\,.$$

$$y_t = f_t(A_t l_t + \varepsilon_t) \qquad \varepsilon_t \sim \mathcal{N}(0,\, \sigma_{2t}^2)\,.$$



شکل ۱: مدل گرافیکی احتمالی Kalman Filter با observation غیر خطی

فیلترینگ در Kalman Filter

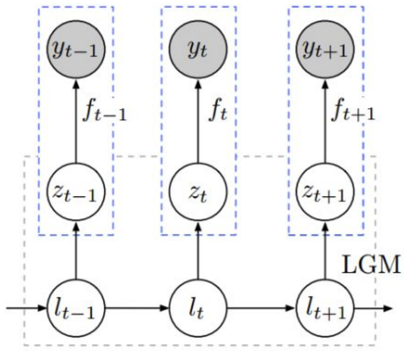


شکل ۱: مدل گرافیکی احتمالی Kalman Filter با observation غیر خطی

یک مسئله مهم در kalman filter مسئله فیلترینگ یعنی $p(l_t|y_{1:t})$ است.

با توجه به اثبات های انجام شده در مقاله میتوان نتیجه گرفت که $p(l_t|y_{1:t}) = p_{LGM}(l_t|z_{1:t})$
بنابراین با توجه به شکل ۱تنها نیاز است تا observationها توسط تابع وارون f به z تبدیل شده و حال یک kalman filter خطی در اختیار داریم که میتوان برای آن روابط بسته فیلترینگ

اثبات



شکل ۱: مدل گرافیگی احتمالی Kalman Filter با observation غیر خطی

برای حل مسئله فیلترینگ دو مرحله وجود دارد. predict و update.

$$predict : p(l_t|y_{1:t-1}) = \int p(l_t|l_{t-1})p(l_{t-1}|y_{1:t-1})dl_{t-1}$$

$$update : p(l_t|y_{1:t}) = \frac{p(y_t|l_t)p(l_t|y_{1:t-1})}{\int p(y_t|l_t)p(l_t|y_{1:t-1})dl_t}$$

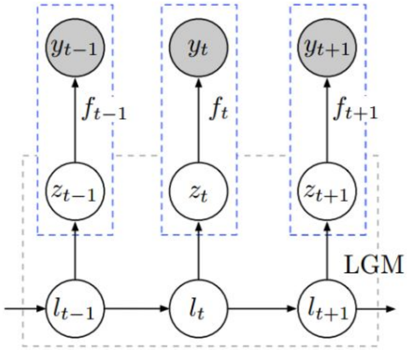
که همان طور که در روابط مشاهده میشود برای محاسبه update که همان جواب مسئله فیلترینگ است از predict استفاده میشود.

اثبات

همچنین با استفاده از روابط normalizing flow میدانیم: $p(y_t|l_t) = p(z_t|l_t)Df_t^{-1}(y_t)$
برای اثبات $p(l_t|y_{1:t}) = p_{LGM}(l_t|z_{1:t})$ از استقرا استفاده کرده ابتدا $p(l_1|y_1)$ را محاسبه میکنیم.

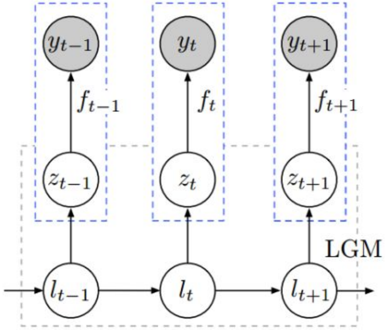
$$p(l_1|y_1) = \frac{p(l_1)p(y_1|l_1)}{\int p(l_1)p(y_1|l_1)dl_1} = \frac{p(l_1)p(z_1|l_1)Df_1^{-1}(y_1)}{\int p(l_1)p(z_1|l_1)Df_1^{-1}(y_1)dl_1}$$
$$= \frac{p(l_1)p(z_1|l_1)}{\int p(l_1)p(z_1|l_1)dl_1} = p_{LGM}(l_1|z_1)$$

حال فرض میکنیم که $p(l_{t-1}|y_{1:t-1}) = p_{LGM}(l_{t-1}|z_{1:t-1})$ است. هم چنین چون $p(l_t|l_{t-1})$ گوسی است بنابراین نتیجه prediction که شامل این دو عبارت است یک توزیع گوسی است که برابر است با $p_{LGM}(l_t|z_{1:t-1})$.



شکل ۱: مدل گرافیگی احتمالی Kalman Filter با observation غیر خطی

اثبات



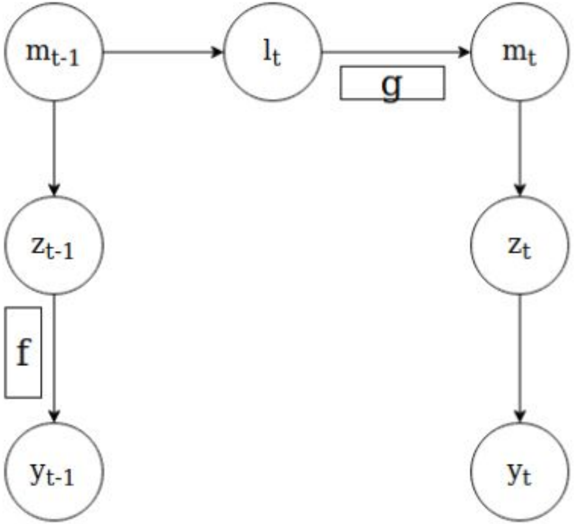
شکل ۱: مدل گرافیکی احتمالی Kalman Filter با observation غیر خطی

در نهایت اگر رابطه update را بسط دهیم:

$$
\begin{aligned}
p(l_t|y_{1:t}) &= \frac{p(y_t|l_t)p(l_t|y_{1:t-1})}{\int p(y_t|l_t)p(l_t|y_{1:t-1})dl_t} \\
&= \frac{Df_t^{-1}(y_t)p(z_t|l_t)p(l_t|y_{1:t-1})}{\int Df_t^{-1}(y_t)p(f_t^{-1}(y_t)|l_t)p(l_t|y_{1:t-1}))dl_t} \\
&= \frac{p(z_t|l_t)p_{LGM}(l_t|z_{1:t-1})}{\int p(z_t|l_t)p_{LGM}(l_t|_{1:t-1}))dl_t} \\
&= p_{LGM}(l_t|z_{1:t})
\end{aligned}
$$

بنابراین با توجه به روابط به دست آمده مدل kalman filterای در اختیار داریم که روابط آن بر اساس مدل خطی است با این تفاوت که observation های غیرخطی y توسط تابع وارون f به z تبدیل میشوند.

مدل kalman filter غیر خطی



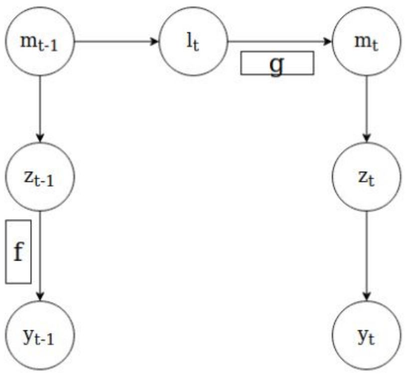
شکل ۲: مدل گرافیکی احتمالی Kalman Filter غیر خطی

$$l_1 \sim \mathcal{N}(\mu_1,\, \sigma_1^2)\,.$$

$$l_t = g_t(F_t l_{t-1} + \epsilon_t) \qquad \epsilon_t \sim \mathcal{N}(0,\, \sigma_{1t}^2)\,.$$

$$y_t = f_t(A_t l_t + \varepsilon_t) \qquad \varepsilon_t \sim \mathcal{N}(0,\, \sigma_{2t}^2)\,.$$

مدل kalman filter غیر خطی



شکل ۲: مدل گرافیکی احتمالی Kalman Filter غیر خطی

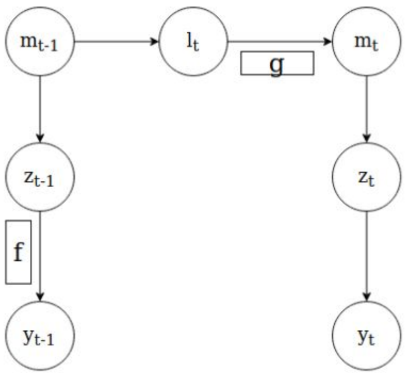
مطابق با مقاله این بار نیز سعی میکنیم تا مسئله فیلترینگ یعنی $p(m_t|y_{1:t})$ را حل کنیم.

هم چنین مطابق با مقاله باید اثبات کنیم:

$$p(m_t|y_{1:t}) \stackrel{?}{=} p_{LGM}(l_t|z_{1:t})$$

بنابراین با توجه به شکل تنها نیاز است تا observationها توسط تابع واورن f به z و هم چنین stateها توسط تابع واورن g به l تبدیل شده و حال یک kalman filter خطی در اختیار داریم که میتوان برای آن روابط بسته فیلترینگ را به دست آورد.

مدل kalman filter غیر خطی



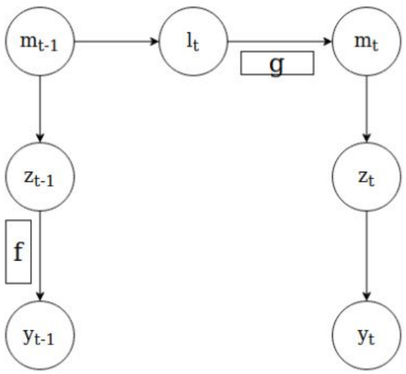
شکل ۲: مدل گرافیکی احتمالی Kalman Filter غیر خطی

روابط predict و update برای مدل جدید مطابق با زیر نوشته می شود.

$$predict : p(m_t|y_{1:t-1}) = \int p(m_t|m_{t-1})p(m_{t-1}|y_{1:t-1})dm_{t-1}$$

$$update : p(m_t|y_{1:t}) = \frac{p(y_t|m_t)p(m_t|y_{1:t-1})}{\int p(y_t|m_t)p(m_t|y_{1:t-1})dm_t}$$

مدل kalman filter غیر خطی



شکل ۲: مدل گرافیکی احتمالی Kalman Filter غیر خطی

همچنین با استفاده از روابط normalizing flow میدانیم:

$$p(y_t|m_t) = p(z_t|m_t)Df_t^{-1}(y_t)$$

$$p(m_t|m_{t-1}) = p(l_t|m_{t-1})Dg_t^{-1}(m_t)$$

مدل kalman filter غیر خطی

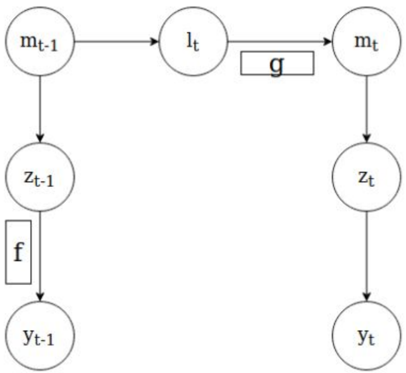
برای اثبات $p(m_t|y_{1:t}) = p_{LGM}(l_t|z_{1:t})$ از استقرا استفاده کرده ابتدا $p(m_1|y_1)$ را محاسبه میکنیم.

$$p(m_1|y_1) = p(l_1|y_1) = \frac{p(l_1)p(y_1|l_1)}{\int p(l_1)p(y_1|l_1)dl_1} = \frac{p(l_1)p(z_1|l_1)Df_1^{-1}(y_1)}{\int p(l_1)p(z_1|l_1)Df_1^{-1}(y_1)dl_1}$$
$$= \frac{p(l_1)p(z_1|l_1)}{\int p(l_1)p(z_1|l_1)dl_1} = p_{LGM}(l_1|z_1)$$

هم چنین $p(l_2|l_1)$ نیز گوسی است.

حال در مرحله استقرا داریم:

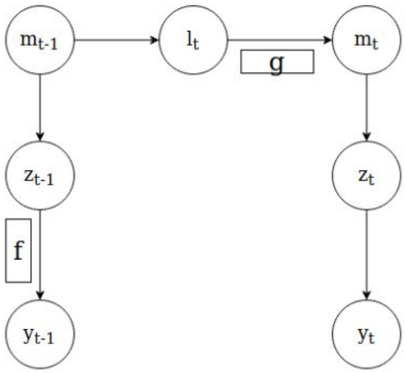
فرض میکنیم که $p(m_{t-1}|y_{1:t-1}) = p_{LGM}(l_{t-1}|z_{1:t-1})$ است.
هم چنین فرض میکنیم $p(l_t|m_{t-1}) = p_{LGM}(l_t|l_{t-1})$ است.



شکل ۲: مدل گرافیکی احتمالی Kalman Filter غیر خطی

$$predict : p(m_t|y_{1:t-1}) = \int p(m_t|m_{t-1})p(m_{t-1}|y_{1:t-1})dm_{t-1}$$

در نهایت اگر رابطه update را بسط دهیم:



شکل ۲: مدل گرافیکی احتمالی Kalman Filter غیر خطی

$$
\begin{aligned}
p(m_t|y_{1:t}) &= \frac{p(y_t|m_t)p(m_t|y_{1:t-1})}{\int p(y_t|m_t)p(m_t|y_{1:t-1})dm_t} \\
&= \frac{Df_t^{-1}(y_t)p(z_t|m_t)p(m_t|y_{1:t-1})}{\int Df_t^{-1}(y_t)p(f_t^{-1}(y_t)|m_t)p(m_t|y_{1:t-1}))dm_t} \\
&= \frac{p(z_t|m_t)p(m_t|y_{1:t-1})}{\int p(f_t^{-1}(y_t)|m_t)p(m_t|y_{1:t-1}))dm_t} \\
&= \frac{p(z_t|m_t)\int p(m_t|m_{t-1})p(m_{t-1}|y_{1:t-1})dm_{t-1}}{\int p(f_t^{-1}(y_t)|m_t)\int p(m_t|m_{t-1})p(m_{t-1}|y_{1:t-1})dm_{t-1})dm_t} \\
&= \frac{p(z_t|m_t)\int p(l_t|m_{t-1})Dg_t^{-1}(m_t)p(m_{t-1}|y_{1:t-1})dm_{t-1}}{\int p(f_t^{-1}(y_t)|m_t)\int p(l_t|m_{t-1})Dg_t^{-1}(m_t)p(m_{t-1}|y_{1:t-1})dm_{t-1})dm_t} \\
&= \frac{p(z_t|m_t)\int p(l_t|m_{t-1})p(m_{t-1}|y_{1:t-1})dm_{t-1}}{\int p(f_t^{-1}(y_t)|m_t)\int p(l_t|m_{t-1})p(m_{t-1}|y_{1:t-1})dm_{t-1})dm_t} \\
&= \frac{p(z_t|m_t)p_{LGM}(l_t|z_{1:t-1})}{\int p(f_t^{-1}(y_t)|m_t)p_{LGM}(l_t|z_{1:t-1})dm_t} \\
&\stackrel{?}{=} \dots \\
&= \frac{p(z_t|l_t)p_{LGM}(l_t|z_{1:t-1})}{\int p(z_t|l_t)p_{LGM}(l_t|z_{1:t-1}))dl_t} \\
&= p_{LGM}(l_t|z_{1:t})
\end{aligned}
$$

Thank you.