Filtr Kalmana

Zygmunt Zawadzki, Mateusz Stachnik

13 maja 2014

Spis treści

1 Wstęp

Dokument ten stanowi zbiór notatek dotyczących filtru Kalmana.

2 Pewne zabawy z rozkładem apriori i aposterori

Zakładamy, że rozkład próbkowy jest to jednowymiarowy rozkład normalny (mamy tylko jedną obserwację), o wartości oczekiwanej μ i znanej wariancji σ^2:

$$p_x(x|\mu) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \tag{1}$$

, natomiast rozkład a priori również jest rozkładem normalnym o wartości oczekiwanej θ i wariancji τ^2.

$$p(\mu) = \frac{1}{\sqrt{2\pi\tau^2}} e^{\frac{-(\mu-\theta)^2}{2\tau^2}} \tag{2}$$

Korzystając z wzoru Bayesa możemy zapisać:

$$p(\mu|x) = \frac{p(x|\mu)p(\mu)}{p(x)} \propto p(x|\mu)p(\mu) \tag{3}$$

, Rozpisując:

$$p(x|\mu)p(\mu) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(x-\mu)^2}{2\sigma^2}}\frac{1}{\sqrt{2\pi\tau^2}}e^{\frac{-(\mu-\theta)^2}{2\tau^2}} \tag{4}$$

, dalej zauważając, że fragmenty $\frac{1}{\sqrt{2\pi\sigma^2}}$ i $\frac{1}{\sqrt{2\pi\tau^2}}$, są niezależne od μ i stałe można dalej zapisać, że:

$$p(x|\mu)p(\mu) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(x-\mu)^2}{2\sigma^2}}\frac{1}{\sqrt{2\pi\tau^2}}e^{\frac{-(\mu-\theta)^2}{2\tau^2}} \propto e^{\frac{-(x-\mu)^2}{2\sigma^2}+\frac{-(\mu-\theta)^2}{2\tau^2}} \tag{5}$$

Teraz należałoby pokazać, że uzyskane równanie jest jądrem znanego rozkładu (a najlepiej normalnego). Gdyby więc udało mi się zapisać 5 w postaci:

$$e^{\frac{-(\mu-\mu_{new})^2}{2\sigma^2_{new}}} \tag{6}$$

,

wtedy natychmiast stwierdziłbym, że jest to jądro rozkładu normalnego i teraz wystarczy tylko dobrać odpowiednią stałą i mam gotowy rozkład a posteriori.

Należy więc dokonać paru prostych przekształceń, jednak w pewnym momencie należy zastosować pewną sztuczkę, dlatego przedstawię wszystko:-). Dla wygody rachunkowej chwilowo pozbywam się e i $-\frac{1}{2}$ i całe równanie zapisuję jako:

$$\begin{aligned}
\frac{(x-\mu)^2}{\sigma^2}+\frac{(\mu-\theta)^2}{\tau^2} &= \frac{x^2+\mu^2-2x\mu}{\sigma^2}+\frac{\theta^2+\mu^2-2\mu\theta}{\tau^2} = \\
&= \frac{(x^2+\mu^2-2x\mu)\tau^2}{\sigma^2\tau^2}+\frac{(\theta^2+\mu^2-2\mu\theta)\sigma^2}{\tau^2\sigma^2} = \\
&= \frac{(x^2+\mu^2-2x\mu)\tau^2+(\theta^2+\mu^2-2\mu\theta)\sigma^2}{\sigma^2\tau^2} = \\
&= \frac{\mu^2(\sigma^2+\tau^2)-2\mu(\tau^2x+\sigma^2\theta)+\tau^2x^2+\sigma^2\theta^2}{\sigma^2\tau^2} = \\
&= \frac{\mu^2(\sigma^2+\tau^2)}{\sigma^2\tau^2}-\frac{2\mu(\tau^2x+\sigma^2\theta)}{\sigma^2\tau^2}+\frac{\tau^2x^2+\sigma^2\theta^2}{\sigma^2\tau^2}
\end{aligned} \tag{7}$$

Gdyby otrzymany wynik wstawić z powrotem do wyrażenia z e, otrzymamy:

$$e^{-\frac{1}{2}(\frac{\mu^2(\sigma^2+\tau^2)}{\sigma^2\tau^2}-\frac{2\mu(\tau^2x+\sigma^2\theta)}{\sigma^2\tau^2}+\frac{\tau^2x^2+\sigma^2\theta^2}{\sigma^2\tau^2})} \tag{8}$$

,

porównując to z rozpisanym wyrażeniem 6:

$$e^{\frac{-(\mu-\mu_{new})^2}{2\sigma^2_{new}}} = e^{-\frac{1}{2}(\frac{\mu^2}{\sigma^2_{new}}-\frac{2\mu\mu_{new}}{\sigma^2_{new}}+\frac{\mu^2_{new}}{\sigma^2_{new}})} \tag{9}$$

, można zauważyć, że są w sumie podobne. W tym momencie, by znaleźć wartość μ_{new} i σ^2_{new}, dobrze jest wykonać pewną sztuczkę. Jeżeli zauważymy, że w równaniach 8 i 9, część $\frac{\tau^2x^2+\sigma^2\theta^2}{\sigma^2\tau^2}$ i $\frac{\mu^2_{new}}{\sigma^2_{new}}$, nie zależą od μ, znaczy to, że można je "wyrzucić" do mitycznej stałej normującej (analogiczny krok jak w przypadku pozbywania się $\frac{1}{\sqrt{2\pi\sigma^2}}$ i $\frac{1}{\sqrt{2\pi\tau^2}}$ z rozkładu aposterori - w zasadzie ciągle na nim pracujemy!). Dla utrzymania uwagi ten krok też przedstawię:

$$
\begin{aligned}
e^{-\frac{1}{2}(\frac{\mu^2(\sigma^2+\tau^2)}{\sigma^2\tau^2}-\frac{2\mu(\tau^2x+\sigma^2\theta)}{\sigma^2\tau^2}+\frac{\tau^2x^2+\sigma^2\theta^2}{\sigma^2\tau^2})} = e^{-\frac{1}{2}(\frac{\mu^2(\sigma^2+\tau^2)}{\sigma^2\tau^2}-\frac{2\mu(\tau^2x+\sigma^2\theta)}{\sigma^2\tau^2})}e^{-\frac{1}{2}\frac{\tau^2x^2+\sigma^2\theta^2}{\sigma^2\tau^2})} \propto \\
\propto e^{-\frac{1}{2}(\frac{\mu^2(\sigma^2+\tau^2)}{\sigma^2\tau^2}-\frac{2\mu(\tau^2x+\sigma^2\theta)}{\sigma^2\tau^2})} \qquad (10) \\
e^{-\frac{1}{2}(\frac{\mu^2}{\sigma^2_{new}}-\frac{2\mu\mu_{new}}{\sigma^2_{new}}+\frac{\mu^2_{new}}{\sigma^2_{new}})} = e^{-\frac{1}{2}(\frac{\mu^2}{\sigma^2_{new}}-\frac{2\mu\mu_{new}}{\sigma^2_{new}})}e^{-\frac{1}{2}\frac{\mu^2_{new}}{\sigma^2_{new}}} \propto e^{-\frac{1}{2}(\frac{\mu^2}{\sigma^2_{new}}-\frac{2\mu\mu_{new}}{\sigma^2_{new}})}
\end{aligned}
$$

,

teraz już praktycznie koniec, teraz wystarczy przyrównać do siebie oba równania i wyliczyć μ_{new} i σ^2_{new}. Z poniższej postaci:

$$
e^{-\frac{1}{2}(\frac{\mu^2}{\sigma^2_{new}}-\frac{2\mu\mu_{new}}{\sigma^2_{new}})} = e^{-\frac{1}{2}(\frac{\mu^2(\sigma^2+\tau^2)}{\sigma^2\tau^2}-\frac{2\mu(\tau^2x+\sigma^2\theta)}{\sigma^2\tau^2})} \qquad (11)
$$

, od razu widać, że $\sigma^2_{new} = \frac{\sigma^2\tau^2}{\sigma^2+\tau^2}$, natomiast przekształcając $\frac{2\mu(\tau^2x+\sigma^2\theta)}{\sigma^2\tau^2}$ w $\frac{2\mu(\tau^2x+\sigma^2\theta)(\sigma^2+\tau^2)}{\sigma^2\tau^2(\sigma^2+\tau^2)}$ i podstawiając σ^2_{new}, otrzymujemy $\frac{2\mu(\tau^2x+\sigma^2\theta)}{\sigma^2_{new}(\sigma^2+\tau^2)}$, łatwo więc zauważyć, że $\mu_{new} = \frac{\tau^2x+\sigma^2\theta}{\sigma^2+\tau^2}$.

Ostatecznie wartość oczekiwana rozkładu aposterori:

$$
\mu_{new} = \frac{\tau^2x+\sigma^2\theta}{\sigma^2+\tau^2} \qquad (12)
$$

, i wariancja:

$$
\sigma^2_{new} = \frac{\sigma^2\tau^2}{\sigma^2+\tau^2} \qquad (13)
$$

Patrząc na wzór 12 można zauważyć, że w zasadzie wartość oczekiwana, rozkładu aposterori przy podanych założeniach, jest średnią ważoną wariancjami rozkładu apriori i aposterori.

Pisząc ostatecznie trochę jednak przesadziłem, gdyż wzory 12 i 13 można przedstawić w trochę innej formie która pozwoli łatwo pewną zależność. W tym przypadku nie będę pisał wszystkich wyprowadzeń, są one dosyć trywialne (jeżeli jednak była by taka potrzeba - proszę mówić):

$$
\mu_{new} = \theta + \frac{\tau^2(x-\theta)}{\sigma^2+\tau^2} = \theta + \frac{\tau^2}{\sigma^2+\tau^2}(x-\theta) \qquad (14)
$$

, i wariancja:

$$
\sigma^2_{new} = \tau^2 - \frac{\tau^4}{\sigma^2+\tau^2} \qquad (15)
$$

Wzór 14 pokazuje, że wartość oczekiwana rozkłądu aposterori, to wartość oczekiwana apriori dodać przeskalowana różnica pomiędzy obserwacją i wartością oczekiwaną apriori. Podana będzie później kluczowa w pokazaniu czym jest filtr Kalmana.

3 Filtr Kalmana

W filtrze Kalmana przyjmuje się, że obserwowany jest pewien proces z_t, który bezpośrednio zależy od procesu ukrytego x_t. Oba procesy można zapisać w postaci:

$$
x_t = Ax_{t-1} + Bu_t + w_t \qquad (16)
$$

,

$$z_t = Hx_t + v_t \tag{17}$$

, gdzie macierz A określa zależność procesu ukrytego w chwili t od stanu w chwili $t-1$, w u_t znajduje się sterowanie procesem (zewnętrzna ingerencja w proces, za chwilę będzie podany pewien intuicyjny przykład), macierz B określa zależność stanu x_t od sterowania, natomiast w_t jest gaussowskim białym szumem o zerowej wartości oczekiwanej i znanej wariancji. Natomiast macierz H jest macierzą przejścia od procesu ukrytego, do obserwowanego.

Na pierwszy rzut oka całe to tak sformułowany model może wydawać się dziwny i nienaturalny - a przez to bezużyteczny! Poniżej jednak pewien przykład:

3.1 Przykład modelu

Jedziemy samochodem. Naszym celem jest szacowanie w każdej chwili prędkości i drogi którą pokonaliśmy od rozpoczęcia podróży. Pojawił się jednak pewien problem natury technicznej - prędkościomierz nie działa! W normalnych warunkach byłoby dosyć niebezpiecznie podróżować takim samochodem, jednak posiadamy GPS, który podaje nam nasze położenie. Dodatkowo uzbrajając się w wiedzę związaną z filtrem Kalmana budujemy stosowny model (dla ułatwienia przyjmiemy również, że hamulce nie działają, przez co nie ma się żadnego wpływu na prędkość:-)):

- x_{1t} - oznacza położenie w chwili t.
- x_{2t} - poszukiwana prędkość w chwili t
- z_t - wskazanie GPS odnośnie pozycji.

Sięgając do wzorów z fizyki znajdziemy, że wzór na drogę w ruchu jednostajnym prostoliniowym to $x_{1t} = x_{1t-1} + (x_{2t-1} \cdot \Delta t)$, tak więc macierze A, B i H będą wyglądały w ten sposób:

$$A = \begin{pmatrix} 1 & \Delta t \\ 0 & 1 \end{pmatrix} \tag{18}$$

,

$$B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \tag{19}$$

$$H' = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{20}$$

3.2 Równania w filtrze Kalmana

Oznaczenia:

- $\hat{x}_{t|t}$- wartość oczekiwana w chwili t, pod warunkiem chwili t
- $\hat{x}_{t|t-1}$- prognoza wartości oczekiwanej na chwilę t, przy znajomości procesu do chwili $t-1$
- $P_{t|t}$- macierz kowariancji dla x_t w chwili t
- $P_{t-1|t-1}$- macierz kowariancji prognozy x_t na chwilę t, pod warunkiem chwili $t-1$

Faza predykcji:

$$\hat{x}_{t|t-1} = A\hat{x}_{t-1|t-1} + Bu_t \tag{21}$$

$$P_{t|t-1} = AP_{t-1|t-1}A' + W \tag{22}$$

Faza uaktualnienia:

$$K_t = (P_{t|t-1}H')(HP_{t|t-1}H' + R)^{-1} \tag{23}$$

,

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t(z_t - H\hat{x}_{t|t-1}) \tag{24}$$

$$P_{t|t} = P_{t|t-1} - K_tHP_{t|t-1} \tag{25}$$

3.3 Wyprowadzenie wzorów - faza predykcji

Poniżej znaduje się wyprowadznie wzorów związanych z fazą predykcjiw filtrze Kalmana - nic trudnego:

Wpierw prognoza wartości oczekiwanej, na okres t, pod warunkiem znajomości wartośći oczekiwanej w czasie $t-1$.

$$\begin{aligned} \hat{x}_{t|t-1} = E(x_t|\hat{x}_{t-1|t-1}) = E(A\hat{x}_{t-1|t-1} + Bu_t + w_t) = \\ = A\hat{x}_{t-1|t-1} + Bu_t + E(w_t) = A\hat{x}_{t-1|t-1} + Bu_t \end{aligned} \tag{26}$$

Komentarz: W powyższym wyprowadzeniu należy jedynie zauważyć, że $\hat{x}_{t-1|t-1}$jest znane, u_t również, a w_t nie zależy od $\hat{x}_{t-1|t-1}$i jego wartość oczekiwana jest równa 0. Jasno widać, że 26 zgadza się z 21.

Teraz będzie weselej - macierz kowariancji dla prognozy x_t ($P_{t|t-1}$):

$$P_{t|t-1} = E[(x_t - \hat{x}_{t|t-1})(x_t - \hat{x}_{t|t-1})'] \tag{27}$$

Na początek dobrze jest policzyć samo:

$$x_t - \hat{x}_{t|t-1} = Ax_t + Bu_t + w_t - A\hat{x}_{t-1|t-1} - Bu_t = A(x_t - \hat{x}_{t-1|t-1}) + w_t \tag{28}$$

Teraz podstawiamy 28 do 27 i otrzymujemy:

$$\begin{aligned} P_{t|t-1} = E[(x_t - \hat{x}_{t|t-1})(x_t - \hat{x}_{t|t-1})'] = E[(A(x_t - \hat{x}_{t-1|t-1}) + w_t)(A(x_t - \hat{x}_{t-1|t-1}) + w_t)'] \\ = AE[(x_t - \hat{x}_{t-1|t-1})(x_t - \hat{x}_{t-1|t-1})']A' + AE[(x_t - \hat{x}_{t-1|t-1})w_t'] + \\ + E[w_t(x_t - \hat{x}_{t-1|t-1})']A' + E(w_tw_t') \end{aligned} \tag{29}$$

Teraz należy zauważyć, że $E[(x_t - \hat{x}_{t-1|t-1})(x_t - \hat{x}_{t-1|t-1})']$ to macierz kowariancji w chwili $t-1$ - czyli $P_{t-1|t-1}$, $E(w_tw_t')$, to macierz kowariancji w_t - czyli W, a kowariancja $E[w_t(x_t - \hat{x}_{t-1|t-1})']$ wynosi 0. Ostatecznie otrzymujemy:

$$P_{t|t-1} = AP_{t-1|t-1}A' + W \tag{30}$$

3.4 Wyprowadzenie wzorów - faza uaktualnienia - intuicyjnie

W tej sekcji zostaną wyprowadzone wzory związane z fazą predykcji - jednak przy pewnym założeniu odnośnie macierzy H - celem takiego podejścia jest pokazanie pewnych związków filtru Kalmana i rekurencyjnego obliczania parametrów rozkładu aposterori pokazanego wcześniej.

Na potrzeby dalszych rozważań przyjmijmy, że macierz H, jest macierzą kwadratową (w ogólności dla filtru Kalmana tak być nie musi) i do tego jest nieosobliwa($det(H) \neq 0$). Wyprowadzone wzory co prawda będą takie same jak w przypadku uchylenia powyższego założenia, jednak proces wyprowadzania będzie zupełnie inny!

Teraz czas na Bayesizm - Na początek przyjmiemy, że w chwili t rozkład apriori dla zmiennej losowej x_t, to $N(\hat{x}_{t|t-1}, P_{t|t-1})$ - takie założenie wydaje się być sensowne - jako rozkład apriori przyjmujemy prognozy parametrów z okresu $t-1$ na okres t. Teraz dobrze byłoby skorzystać ze wzorów 14 i 15, problemem jest jednak to, że x_t jest nieobserwowalne - tzn. nie ma nadziei, że pojawi się nowa obserwacja. Mamy jednak dostęp realizacji z_t. Można więc spróbować na podstawie wzorów 14 i 15 wyznaczyć rozkład aposterori dla zmiennej losowej Hx_t, a potem powrócić do x_t. Rozkład apriori dla Hx_t to $N(H\hat{x}_{t|t-1}, HP_{t|t-1}H')$. Ze wzoru 14 mamy:

$$H\hat{x}_{t|t} = H\hat{x}_{t|t-1} + \frac{HP_{t|t-1}H'}{HP_{t|t-1}H' + R}(z_t - H\hat{x}_{t|t-1}) \tag{31}$$

, mnożąc lewostronnie przez H^{-1} (właśnie po to było założenie o nieosobliwości macierzy H), otrzymujemy:

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + (P_{t|t-1}H')(HP_{t|t-1}H' + R)^{-1}(z_t - H\hat{x}_{t|t-1}) \tag{32}$$

Natomiast dla kowariancji:

$$HP_{t|t}H' = HP_{t|t-1}H' - (HP_{t|t-1}H')(HP_{t|t-1}H' + R)^{-1}(HP_{t|t-1}H')' \tag{33}$$

, i analogicznie jak w 31 przez H^{-1} i $(H')^{-1}$, otrzymując:

$$P_{t|t} = P_{t|t-1} - (P_{t|t-1}H')(HP_{t|t-1}H' + R)^{-1}HP_{t|t-1} \tag{34}$$

, dla wygody zauważamy, że w obu wypadkach 32 i 34 występuje ten sam fragment postaci: $(P_{t|t-1}H')(HP_{t|t-1}H' + R)^{-1}$ - dobrze nazwać go sobie jakoś nazwać - np K_t:

$$K_t = (P_{t|t-1}H')(HP_{t|t-1}H' + R)^{-1} \tag{35}$$

, i po podstawieniu do 32 i 34, otrzymujemy:

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t(z_t - H\hat{x}_{t|t-1}) \tag{36}$$

$$P_{t|t} = P_{t|t-1} - K_tHP_{t|t-1} \tag{37}$$

, teraz uważny obserwator porównując wzory 35, 36, 37 ze wzorami 23, 24 i 25, zauważy, że są identyczne!

Niestety trzeba pamiętać o przyjętym założeniu, że H jest macierzą nieosobliwą. W następnej sekcji powyższe wzory zostaną wyprowadzone w trochę inny sposób.

3.5 Wyprowadzanie wzorów - uaktualnianie

Powyżej wzory związane z filtrem Kalmana zostały wyprowadzone przy założeniu, że macierz H jest macierzą nieosobliwą - jest to bardzo krępujące, gdyż w większości zastosować H nie jest nawet kwadratowa (chociażby w przykładzie w 3.1)!

By pozbyć się tego założenia, trzeba podejść do problemu w inny sposób. Zacznijmy od intuicji - x_t jest nieobserwowane, przez co nie ma szans na bezpośrednie zastosowanie formuły 14. Jeżeli H nie jest kwadratowa, to nie będą możliwe do wykonania kroki przedstawione w poprzedniej sekcji. Mamy więc do dyspozycji realizacje z_t i predykcję wartości oczekiwanej z poprzedniego okresu $\hat{x}_{t|t-1}$. Trzeba zastosować podejście w pewnym sensie inżynieryjne. Przyjmijmy, że im większe różnica pomiędzy wartościami które możemy wyznaczyć $(z_t - H\hat{x}_{t|t-1})$, tym zmiana $\hat{x}_{t|t}$, powinna być większa. Jako, że różnica $z_t - H\hat{x}_{t|t-1}$ nie koniecznie musi być w tym samym wymiarze, co $\hat{x}_{t|t}$, należy wykorzystać jakąś macierz - dowolnej postaci!!! - która sprawi, że wymiary pomiędzy $\hat{x}_{t|t}$ i $z_t - H\hat{x}_{t|t-1}$ będą się zgadzać. Taką macierz możnaby nazwać na przykład Z_t. Otrzymujemy więc:

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + Z_t(z_t - H\hat{x}_{t|t-1}) \tag{38}$$

. Powyższe równanie wygląda znajomo - tylko zamiast K_t występuje Z_t.

W tej chwili Z_t jest dowolną macierzą, której jedyną własnością, jest to, że wymiary macierzy w równaniu 38 zgadzają się ze sobą. Należałoby przyjąć jakieś kryterium, względem którego możnaby wyznaczyć najbardziej optymalną postać Z_t. Tym kryterium może być, np. minimalizacja wartości oczekiwanej kwadratów reszt - czyli:

$$E(||x_t - \hat{x}_{t|t}||^2) \to min. \tag{39}$$

. Tak przedstawiony problem, jest równoważny minimalizacji śladu macierzy kowariancji $Cov(x_t - \hat{x}_{t|t}) = P_{t|t}$. Należy zwrócić uwagę, że macierz $P_{t|t}$- na razie nie ma nic wspólnego z macierzą wyznaczoną w poprzednim podrozdziale -jednak takie oznaczenie jest bardzo wygodne. Można powiedzieć, że należy zapomnieć o wszystkim z 3.4 (natomiast wszystkie wzory z 3.3 pozostają w mocy).

$$\begin{aligned} P_{t|t} = Cov(x_t - \hat{x}_{t|t}) = Cov[x_t - \hat{x}_{t|t-1} - Z_t(z_t - H\hat{x}_{t|t-1})] = \\ = Cov[x_t - \hat{x}_{t|t-1} - Z_t z_t + Z_t H\hat{x}_{t|t-1}] = Cov[x_t - \hat{x}_{t|t-1} - Z_t H x_t - Z_t v_t + Z_t H\hat{x}_{t|t-1}] = \\ = Cov[I(x_t - \hat{x}_{t|t-1}) - Z_t H(x_t - \hat{x}_{t|t-1}) - Z_t v_t] = Cov[(I - Z_t H)(x_t - \hat{x}_{t|t-1}) - Z_t v_t] \end{aligned} \tag{40}$$

, Powyższe przejścia wyglądają źle, ale nie są trudne, następuje podstawienie wzoru 38, następnie, za z_t, podstawione zostaje 17. Potem następują zwykłe przekształcenia algebraiczne. By uprościć obliczenia należy zauważyć, że kowariancja pomiędzy v_t, a resztą wynosi 0 i wtedy możemy zapisać:

$$\begin{aligned} P_{t|t} = Cov[(I - Z_t H)(x_t - \hat{x}_{t|t-1})] + Cov[Z_t v_t] = \\ = (I - Z_t H)Cov[(x_t - \hat{x}_{t|t-1})](I - Z_t H)' + Z_t Cov[v_t]Z_t' = \\ = (I - Z_t H)P_{t|t-1}(I - Z_t H)' + Z_t R Z_t' \end{aligned} \tag{41}$$

Tę postać należy zapamiętać:

$$P_{t|t} = (I - Z_t H)P_{t|t-1}(I - Z_t H)' + Z_t R Z_t' \tag{42}$$

, poniżej jeszcze wersja rozpisana, która przyda się już za chwilę:

$$\begin{aligned} P_{t|t} = P_{t|t-1} - P_{t|t-1}Z_t'H' - Z_tHP_{t|t-1} + Z_tHP_{t|t-1}H'Z_t' + Z_tRZ_t' = \\ = P_{t|t-1} - P_{t|t-1}Z_t'H' - Z_tHP_{t|t-1} + Z_t(HP_{t|t-1}H' + R)Z_t' \end{aligned} \tag{43}$$

Teraz szukamy, Z_t minimalizującego ślad macierz $P_{t|t}$:

$$\begin{aligned} \frac{dP_{t|t}}{dZ_t} = -2P_{t|t-1}H' + 2Z_t(HP_{t|t-1}H' + R) = 0 \\ Z_t(HP_{t|t-1}H' + R) = P_{t|t-1}H' \\ Z_t = P_{t|t-1}H'(HP_{t|t-1}H' + R)^{-1} \end{aligned} \tag{44}$$

, Porównując ten wzór z 23, łatwo zauważyć, że $Z_t = K_t$, czyli K_t w filtrze Kalmana minimalizuje sumę kwadratów reszt.

4 Prosta implementacja filtra Kalmana w R

```
# Wzory na faze predyckji:
calcXt1 = function(A, xt, B = NULL, u) {
    if (is.null(B))
        return(A %*% xt)
    return(A %*% xt + B %*% u)
}

# Faza uaktualniania:
calcPt1 = function(A, Pt, Q) A %*% Pt %*% t(A) + Q
kalmanGain = function(Pt1, H, R) Pt1 %*% t(H) %*% solve(H %*% Pt1 %*% t(H) +
    R)
calcXt = function(xt1, K, z, H) xt1 + K %*% (z - H %*% xt1)
calcPt = function(Pt1, K, H) Pt1 - K %*% H %*% Pt1
```