

MI Metody Identyfikacji

wykład #10a

- 1. Filtr Kalmana
- 2. Obserwator stanu
- 3. Filtr Bayesa
- 4. Rozszerzony filtr Kalmana

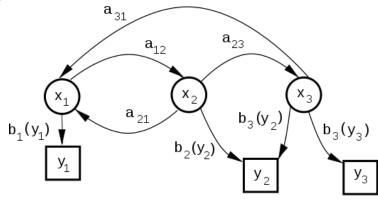
Estymacja stanu

- Zadanie polega na estymacji stanu procesu $\hat{x}(k)$ na podstawie wejść u(l) i wyjść y(l) aż do chwili j
- W zależności od doboru chwili j i k mamy różne zadania
 - k > j predykcja na n-kroków (n=k-j)
 - k = j zadanie filtrowania
 - k < j zadanie wygładzania
- Rozważymy zadanie predykcji na jeden krok naprzód
- Filtr Kalmana (1960 *A new approach to linear filtering and prediction problem*) jest zdefiniowany całkowicie w dziedzinie czasu



Porównanie łańcuchów Markowa i filtru Kalmana

- Łańcuch markowa (HMM Hidden Markov Models)
 - Proces stochastyczny
 - Dyskretne zmienne stanu
 - Służy do modelowania sekwencji zdarzeń



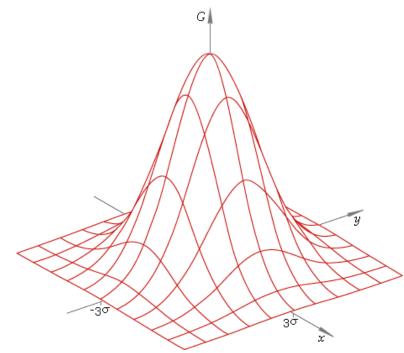
- Filtry Kalmana
 - Zmienne ciągłe w czasie o rozkładzie normalnym
 - Wykorzystywany do modelowania zaszumionych ciągłych obserwacji
 - W podstawowej definicji stacjonarny i liniowy ale możliwe rozszerzenie na procesy

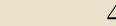
nieliniowe i niestacjonarne



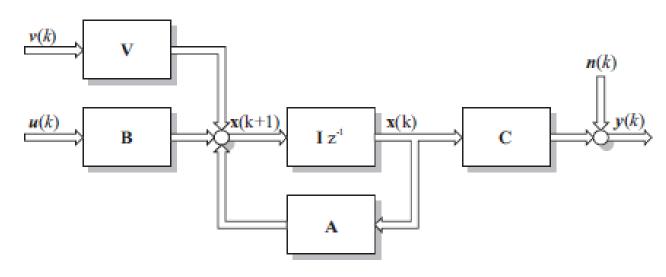
Filtr Kalmana

- Algorytm rekurencyjnego wyznaczania minimalno-wariancyjnej estymaty wektora stanu modelu liniowego dyskretnego układu dynamicznego na podstawie pomiarów wyjścia tego układu.
- Przyjmuje się założenie, że zarówno wejście u(k), jak i stan układu x(k) są zmiennymi o rozkładzie gaussowskim i zerowej wartości średniej.





Opis w postaci równań stanu



$$x(k+1) = A \cdot x(k) + B \cdot u(k) + V \cdot v(k)$$
$$y(k) = C \cdot x(k) + n(k)$$

v(k) i n(k) są nieskorelowanymi białymi szumami o zerowej wartościach średnich i kowariancjach

$$E\{v(k)v^{T}(k)\} = M$$

$$E\{n(k)n^{T}(k)\} = N$$



Filtr Kalmana

- Poszukujemy optymalnego filtru zmiennych stanu o najmniejszym możliwym błędzie
- Zatem staramy się minimalizować błąd kwadratowy

$$V = E\{\|\hat{x}(k+1) - x(k+1)\|_{2}^{2}\}$$

= $E\{(\hat{x}(k+1) - x(k+1))^{T}(\hat{x}(k+1) - x(k+1))\}$

- Idea (predyktor / korektor):
 - Prognozujemy stan na jeden krok naprzód na chwilę k+1
 - Korygujemy prognozę uwzględniając dostępny pomiar wyjścia w chwili k+1 czyli y(k+1)



Filtr Kalmana

- Założenia:
 - x(k) prawdziwy stan w chwili k
 - $\hat{x}(k+1|k)$ predykcja stanu na podstawie pomiarów do chwili k
 - $\hat{x}(k+1|k+1)$ predykcja stanu na podstawie pomiarów do chwili k+1
 - P(k) macierz kowariancji stanów

$$P(k) = E\left\{ \left(\hat{x}(k) - x(k)\right) \left(\hat{x}(k) - x(k)\right)^T \right\}$$



Filtr Kalmana – predykcja

Wyznaczamy predykcję

$$x(k+1) = A \cdot x(k) + B \cdot u(k) + V \cdot v(k)$$

- Zakłócenie v(k) nie jest znane.
- Ma zerową wartość oczekiwaną.
- Tym samym estymata stanu ma postać:

$$\hat{x}(k+1|k) = A \cdot \hat{x}(k) + B \cdot u(k)$$
 wektor obserwacji





Filtr Kalmana – predykcja

• Wyznaczamy nową macierz kowariancji $P^-(k+1)$

$$P^{-}(k+1) = \mathbb{E}\left\{ \left(\hat{x}(k+1|k) - x(k+1) \right) \left(\hat{x}(k+1|k) - x(k+1) \right)^{\mathsf{T}} \right\}$$

$$= \mathbb{E}\left\{ \left(A\hat{x}(k) - Ax(k) + V\nu(k) \right) \left(A\hat{x}(k) - Ax(k) + V\nu(k) \right) \right\}$$

$$= A\mathbb{E}\left\{ \left(\hat{x}(k) - x(k) \right) \left(\hat{x}(k) - x(k) \right)^{\mathsf{T}} \right\} A^{\mathsf{T}}$$

$$+ A\mathbb{E}\left\{ \left(\hat{x}(k) - x(k) \right) \nu^{\mathsf{T}} \right\} V^{\mathsf{T}} \qquad \text{Pamietan prognoza}$$

$$+ V\mathbb{E}\left\{ \nu(k) \left(\hat{x}(k) - x(k) \right)^{\mathsf{T}} \right\} A^{\mathsf{T}} \qquad \text{prognoza}$$

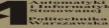
$$+ V\mathbb{E}\left\{ \nu(k) \nu^{\mathsf{T}}(k) \right\} V^{\mathsf{T}} \qquad \text{ma zerow}$$

Pamiętamy, że zarówno stan jak i jego prognoza nie są skorelowane z zakłóceniem jak również zakłócenie ma zerową wartość średnią

Otrzymując ostatecznie (wciąż przed korekcją):

$$P^{-}(k+1) = A \cdot P(k) \cdot A^{T} + V \cdot M \cdot V^{T}$$
 residuum macierzy kowariancji





Filtr Kalmana – korekcja

- Pojawia się nowy pomiar wyjścia y(k+1)
- Dokonujemy korekcji:

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1) \cdot (y(k+1) - C \cdot \hat{x}(k+1|k))$$

 Wybór wzmocnień sprzężenia zwrotnego określa czy predykcja jest wyznaczana na podstawie wewnętrznego modelu czy też pomiaru

$$y(k+1) - C \cdot \hat{x}(k+1|k)$$
 - błąd obserwacji, tzw. innowacja



Filtr Kalmana – korekcja

- Jak wyznaczyć macierz K(k+1)?
- Modyfikujemy prognozę stanu

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1) \cdot \left(C \cdot x(k+1) + n(k+1) - C \cdot \hat{x}(k+1|k)\right)$$

Co prowadzi do postaci macierzy kowariancji

$$P(k+1) = E\left\{ \left(\hat{x}(k+1|k+1) - x(k+1) \right) \left(\hat{x}(k+1|k+1) - x(k+1) \right)^T \right\}$$



Filtr Kalmana – korekcja

optymalne wzmocnienie filtru Kalmana

$$K(k+1) = P^{-}(k+1)C^{T}(C \cdot P^{-}(k+1) \cdot C^{T} + N)^{-1}$$

poprawiony wektor stanu

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1) \cdot (y(k+1) - C \cdot \hat{x}(k+1|k))$$

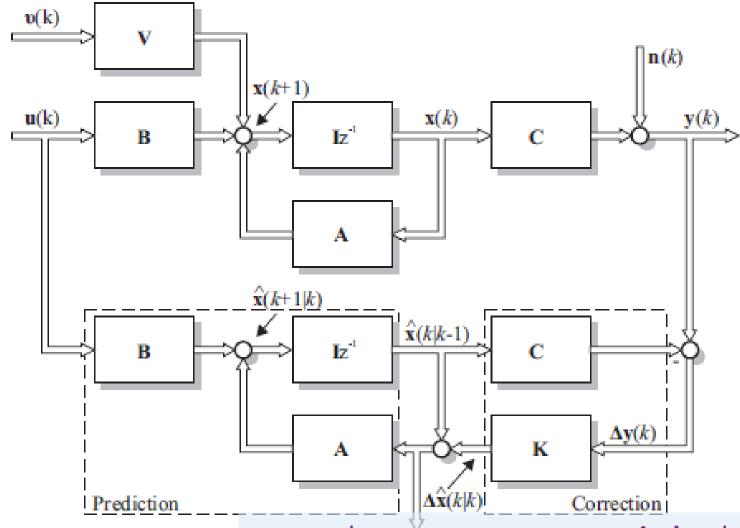
poprawiona macierz kowariancji

$$P(k+1) = [I - K(k+1) \cdot C] \cdot P^{-}(k+1)$$

Wzór ten jest prawdziwy tylko w przypadku, gdy wzmocnienie filtru Kalmana jest optymalne



Filtr Kalmana – schemat blokowy



optymalna estymata stanu na jeden krok naprzód

Niezmienniczość

• Jeśli nasz model jest dokładny, zaś wartości startowe x(0) i P(0) precyzyjnie opisują wektor stanu wówczas następujące wielkości pozostają zachowane (wartości oczekiwane są równe zero):

$$E\{x(k) - \hat{x}(k|k)\} = E\{x(k) - \hat{x}(k|k-1)\} = 0$$
$$E\{\tilde{y}(k)\} = 0$$

 Ponadto macierze kowariancji dokładnie odzwierciedlają macierze kowariancji błędów.



Filtr Kalmana w stanie ustalonym

- W ogólnej formie filtr Kalmana wiąże się z dużym kosztem obliczeniowym
- Dla procesów stacjonarnych można pokazać, że zarówno P(k) jak i K(k) dążą do wartości stałych, dla $k{\longrightarrow}\infty$
- Wartości te można wyznaczyć a priori co znakomicie przyspieszy obliczenia
- Predykcja

$$\hat{x}(k+1|k) = A \cdot \hat{x}(k) + B \cdot u(k)$$

Korekcja

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + \overline{K} \cdot \left(y(k+1) - C \cdot \hat{x}(k+1|k) \right)$$

$$\overline{K} = \overline{P}^- \cdot C \cdot [C \cdot \overline{P}^- \cdot C^T + N]^{-1}$$



Filtr Kalmana - rozszerzenia

- Filtr Kalmana można rozszerzyć na procesy niestacjonarne
- Rozszerzony filtr Kalmana (EKL Extended Kalman Filter) procesy nieliniowe (wykorzystanie linearyzacji wokół punktu pracy)
- Rozszerzony filtr Kalmana do estymacji parametrów
- Istnieją rozszerzania filtru Kalmana przy założeniach nie-Gaussowskich



Podsumowanie

- Filtr Kalmana odpowiada metodzie wariacyjnej dla układu liniowego
- Dla układów z niewielką nieliniowością stosuje się rozszerzony filtr Kalmana. Jednak wymaga to użycia modelu sprzężonego oraz operatorów stycznych



Prosty przykład

- Estymujemy losowa stałą → "voltage" odczytywaną ze źródła
- Posiada ona stałą wartość **a**V [V] i nie ma sygnału wejściowego u_k .
- Odchylenie standardowe szumu pomiarowego wynosi 0.1 [V].
- Problem jest jednowymiarowy:
 macierze A i H (poprzednio oznaczana jako C) są stałymi 1
- Zakładamy jednostkową początkową kowariancję $P_{\theta} = 1$ a początkowy stan zerowy $x_{\theta} = 0$.



Prosty przykład – cd.

$$x_k = Ax_{k-1} + Bu_k + w_k$$
$$= x_{k-1} + w_k$$

$$z_k = Hx_k + v_k$$
$$= x_k + v_k$$

czas	1	2	3	4	5	6	7	8	9	10
wartość	0.39	0.50	0.48	0.29	0.25	0.32	0.34	0.48	0.41	0.45





Time Update (prediction)	Measurement Update (correction)
$\hat{x}_k^- = \hat{x}_{k-1}$	$K_k = \frac{P_k^-}{}$
$P_k^- = P_{k-1}$	$^{\kappa}$ $P_{k}^{-} + R$
	$\hat{x}_k = \hat{x}_k + K_k(z_k - \hat{x}_k)$
	$P_k = (1 - K_k) P_k$



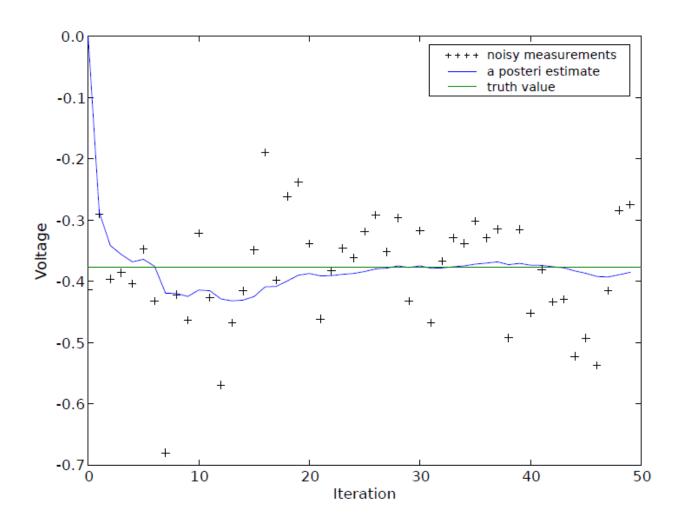


k	1	2	3	4	5	6	7	8	9	10
z_k	0.390	0.500	0.480	0.290	0.250	0.320	0.340	0.480	0.410	0.450
\hat{x}_{k-1}	0	0.355	0.424	0.442	0.405	0.375	0.365	0.362	0.377	0.380
P_{k}^{-}	1	0.091	0.048	0.032	0.024	0.020	0.016	0.014	0.012	0.011
Time Update	$\hat{x}_k = \hat{x}_{k-1} = 0$	$\hat{x}_{k}^{-} = 0.355$								
	$P_{k} = P_{k-1} = 1$	$P_{k}^{-} = 0.091$								
	K _k = 1 / (1 0.1)	K _k = 0.091 / (0.091 0.1)								
	= 0.909	= 0.476								
Measurement	$\hat{x}_k = 0.909 (0.390 - 0)$	$\hat{x}_k = 0.355 \ 0.476 \ (0.500 - 0.355)$								
Update	= 0.35	= 0.424								
	$P_k = (1 - 0.909) . 1$	$P_k = (1 - 0.476) \cdot 0.091$								
	= 0.091	= 0.048								
\hat{x}_k	0.355	0.424	0.442	0.405	0.375	0.365	0.362	0.377	0.380	0.387
P_{k}	0.091	0.048	0.032	0.024	0.020	0.016	0.014	0.012	0.011	0.010





Prosty przykład – cd.





Projektowanie filtru Kalmana

- Wybór wektora stanu
- Wybór zakłóceń
- Wybór zmiennych mierzonych
- Opis szumu pomiarowego
- Wybór macierzy kowariancji
- Wyznaczenie wektora wzmocnień
- Inicjalizacja
- Kwestia brakujących danych
- Testowanie
- Metody obliczeniowe



Zastosowania filtru Kalmana

- Popularne zastosowania
 - nawigacja, naprowadzanie, śledzenie radarowe, przeszukiwanie sonarowe, wyznaczanie orbit satelit, prognozowanie wartości akcji, lądowanie na Księżycu, żyroskopy lotnicze, etc.
- Rozszerzony Filtr Kalmana (EKF) uwzględnia nieliniowości w rozkładach
 - Prognozowanie w otoczeniu zlinearyzowanym otoczeniu punktu pracy
 - Dobry dla gładkich przewidywalnych procesów
- Przełączane filtry Kalmana: różne filtry, każdy z innym modelem
 - Predykcja w postaci ważonej sumy prognoz



Zastosowania filtru Kalmana

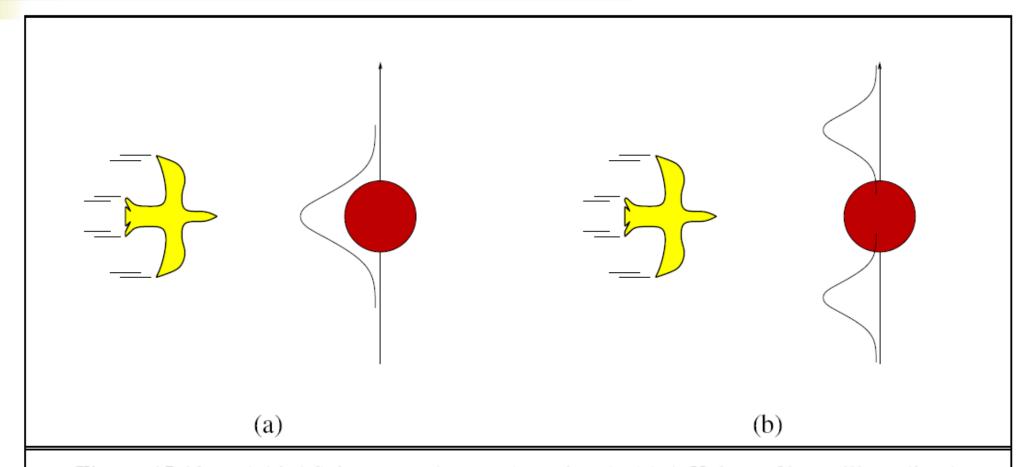


Figure 15.10 A bird flying toward a tree (top views). (a) A Kalman filter will predict the location of the bird using a single Gaussian centered on the obstacle. (b) A more realistic model allows for the bird's evasive action, predicting that it will fly to one side or the other.

26

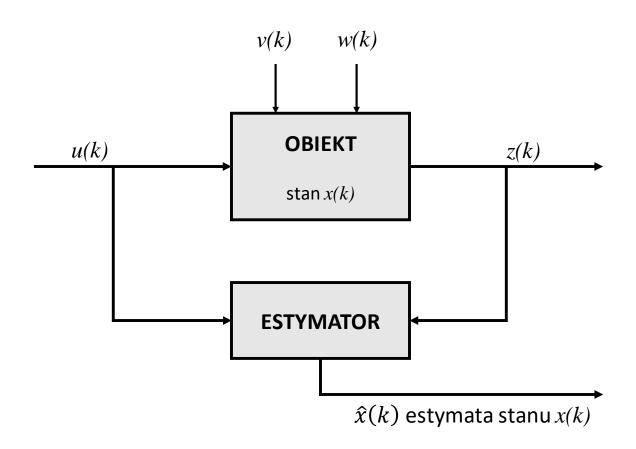
Obserwator stanu

- Filtra Kalmana jest przykładem obserwatora stanu
 - u(k): znane wejście
 - z(k): mierzone wyjście
 - v(k): zakłócenie obiektowe
 - w(k): szum pomiarowy
 - x(k): stan obiektu
- Nieliniowy obiekt opisany dyskretnymi równaniami stanu

$$x(k) = q_{k-1}(x(k-1), u(k-1), v(k-1))$$

$$z(k) = h_k(x(k), w(k))$$

gdzie x(o), $\{v(.)\}$, $\{w(.)\}$ mają opis stochastyczny



2018 zima

Obserwator stanu

- Cel → rozwiązania rekurencyjne
 - Wyznaczanie estymaty stanu $\hat{x}(k)$ na podstawie informacji dostępnej w chwili k.
 - Obliczamy $\hat{x}(k)$ na podstawie $\hat{x}(k-1)$, u(k-1), z(k) oraz wiedzy o modelu (model dynamiczny procesu oraz model stochastyczny zakłóceń)
 - Brak konieczności śledzenia pełnej historii wejść i wyjść $\{u(.)\}, \{z(.)\}$

Zastosowania:

- Trzeba pamiętać, że estymacja jest dualna względem regulacji.
 - Sprzężenie zwrotne na podstawie stanu x(k) wyznacza wejście u(k).
 - Tym samym regulacja ze sprzężeniem zwrotnym jako taka, gdy nie mamy dostępu do stanu.
- Estymacja bez zamykania pętli, czyli estymacja zmiennych sama w sobie:
 - monitorowanie,
 - analiza awaryjności,
 - lokalizacja,
 - diagnostyka w medycynie.



2018 zima

27

Algorytmy

- Zatem mówimy o algorytmach typu:
 - Filtrowanie Bayesowskie
 - Systemy liniowe i rozkłady Gaussowskie filtr Kalmana
 - Systemy nieliniowe i rozkłady Gaussowskie rozszerzony filtr Kalmana
 - Systemy nieliniowe i rozkłady nie-Gaussowskie Particle Filter



Twierdzenie Bayesa

- Prawdopodobieństwo warunkowe p(a/b), jest wynikiem zdarzeń połączonych $p(a/b) = p(a \cap b) / p(b)$
- Prawdopodobieństwo iloczynu zdarzeń: $p(a \cap b) = p(a)p(b/a)$
- Znając prawdopodobieństwo warunkowe i iloczynu zdarzeń możemy wywnioskować, iż iloczyn jest przemienny: $p(a \cap b) = p(b \cap a)$
- Z twierdzenia o prawdopodobieństwie iloczynu zdarzeń mamy: $p(a \cap b) = p(a)p(b/a)$
- Podstawiając powyższe do wzoru $p(a \cap b) = p(b \cap a)$ otrzymujemy: $p(b) \ p(a/b) = p(a) \ p(b/a)$
- co prowadzi do wzoru Bayesa:

$$p(a|b) = \frac{p(a) \cdot p(b|a)}{p(b)}$$



Twierdzenie Bayesa

$$p(a|b) = \frac{p(a) \cdot p(b|a)}{p(b)}$$

- Jeśli skutek A nastąpił w wyniku zajścia jednej z przyczyn $B_1,\,B_2,\,...\,B_n$, to prawdopodobieństwo tego, że B_i była przyczyną zajścia A wyraża się wzorem Bayesa. Prawdopodobieństwo $p(B_i)$ czasami nazywamy prawdopodobieństwem a priori, natomiast prawdopodobieństwo $p(B_i/A)$ nazywamy prawdopodobieństwem a posteriori, gdyż podaje ono szansę zajścia B_i po zaobserwowaniu zajścia zdarzenia A.
 - A nieznana zmienna (stan)
 - p(a) ocena zmiennej *a priori*
 - B obserwacja powiązana ze stanem (pomiar)
 - p(B|A) obserwacja modelu: jakie jest prawdopodobieństwo obserwacji B dla danego stanu
 - p(A|B) a posteriori ocena stanu, po uwzględnieniu obserwacji
 - P(B) prawdopodobieństwo obserwacji (nie zależy od A)



- <u>Przypadek 1:</u> Rodzina ma dwoje dzieci. Jakie jest prawdopodobieństwo, że oboje to dziewczynki?
- Odpowiedź: $\frac{1}{2}$



- Przypadek 2: Rodzina ma dwoje dzieci. Jeśli jedna to dziewczynka, to jakie jest prawdopodobieństwo, że druga jest dziewczynką?
- Odpowiedź:
- Oznaczamy: $a = \begin{cases} 1: & nie \ ma \ chłopców \\ 0: & jest \ chłopiec \end{cases}$ $b = \begin{cases} 1: & nie \ ma \ dziewczynek \\ 0: & jest \ dziewczynka \end{cases}$
- Zatem: $p_a(1) = \frac{1}{4}$, $p_a(0) = \frac{3}{4}$, $p_b(1) = \frac{1}{4}$, $p_b(0) = \frac{3}{4}$
- I otrzymujemy:

$$p_{a|b}(1|0) = \frac{p_a(1) \cdot pb_{a}(0|1)}{p_b(0)} = \frac{1/4 \cdot 1}{3/4} = 1/3$$



- <u>Przypadek 3:</u> Lekarze mieli podać prawdopodobieństwo, że kobieta bez symptomów, pomiędzy 40 a 50 rokiem życia, która ma pozytywny mammograf ma raka piersi?
 - Dane:

7% mammografów przekłamuje pozytywny wynik, 10% mammografów przekłamuje wynik negatywny, prawdopodobieństwo raka w grupie 40-50 wynosi 0.8%. Niemcy: jedna trzecia lekarzy podała 90%, mediana wynosiła 70% USA: 95% lekarzy podało, że około 93%

• Wynik:

• Oznaczamy: $a = \begin{cases} 1: & pacjent \ nie \ ma \ raka \\ 0: & pacjent \ ma \ raka \end{cases}$ $b = \begin{cases} 1: & mammograf \ daje \ wynik \ negatywny \\ 0: & mammograf \ daje \ wynik \ pozytywny \end{cases}$

Zatem szukamy:

$$p_{a|b}(0|0) = \frac{p_a(0) \cdot pb_{a}(0|0)}{p_b(0)}$$



- Wyznaczamy:
 - $p_{bla}(0|0) = 0.9$ (10% mammografów przekłamuje wynik negatywny)
 - $p_b(0) = p_{b|a}(0|0) p_a(0) + p_{b|a}(0|1) p_a(1) = 0.90 \times 0.008 + 0.07 \times 0.092$

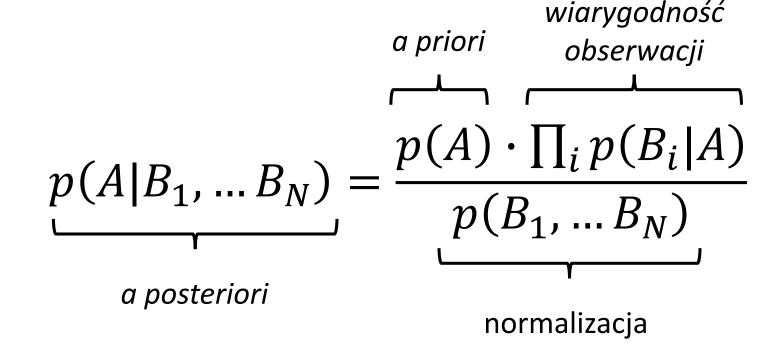
•
$$p_{a|b}(0|0) = \frac{0.008 \cdot 0.9}{0.9 \times 0.008 + 0.07 \times 0.992} \approx 0.094$$

(większość pozytywnych wyników wynika z błędów pozytywnej oceny)



Generalizacja na wiele obserwacji

- Mamy N obserwacji: B₁, ..., B_N
- Można założyć niezależność, tzn. $p(B_1, ..., B_N|A) = p(B_1|A)... p(B_N|A)$ Założenie, że szum jest niezależny
- Wtedy





Śledzenie Bayesowskie

- Algorytm ma dwa kroki:
 - 1. Aktualizacja *a priori*, tzn. estymacja stanu na podstawie modelu
 - 2. Uwzględnienie pomiaru (wykorzystując twierdzenie Bayesa) a posteriori
- x(k) stan
- z(k) obserwacje
- Model (k = 1, 2, ...) $x(k) = q_{k-1}(x(k-1), v(k-1))$ $z(k) = h_k(x(k), w(k))$
- gdzie x(0), $\{v(.)\}$, $\{w(.)\}$ są niezależne o nieznanych funkcjach gęstości prawdopodobieństwa





Śledzenie Bayesowskie – cel

- Niech z(1:k) oznacza zbiór $\{z(1), ... z(k)\}$. Chcemy wyznaczyć efektywnie f(x(k)/z(1:k)), co oznacza $f_{x(k)|z(1:k)}\bar{x}(k)|\bar{z}(1:k)$, gdzie:
 - *x*(*k*) jest zmienna losową
 - $\bar{x}(k)$ jest pewną wartością jaką x(k) może przyjąć
 - *z*(1:*k*) jest zmienna losową
 - $\bar{z}(1:k)$ jest obserwacją z(1:k)



Algorytm rekurencyjny

• Po pierwsze:

$$f(x(k)|z(1:k-1)) = \sum_{x(k-1) \in \mathcal{X}} f(x(k)|z(1:k-1), x(k-1)) \underbrace{f(x(k-1)|z(1:k-1))}_{\text{zakładamy, że znane}}.$$

- x(k) oraz z(1:k-1) są warunkowo niezależne wg
 - $-x(k) = q_{k-1}(x(k-1), v(k-1)),$ funkcja jedynie v(k-1)
 - $z(k-1) = h_{k-1}(x(k-1), w(k-1))$

$$z(k-2) = h_{k-2}(x(k-2), w(k-2)), x(k-2) = q_{k-3}(x(k-3), v(k-3)), \text{ etc.}$$

Tym samym:

$$z(1:k-1) = \text{FUNCTION}\left(x(k-1), \underbrace{v(1:k-3), w(1:k-1), x(0)}_{\text{niezależne od } v(k-1)}\right)$$

• Otrzymujemy: f(x(k)|z(1:k-1), x(k-1)) = f(x(k)|x(k-1))



Algorytm rekurencyjny cd.

• Aktualizacja *a priori*:

$$f(x(k)|z(1:k-1)) = \sum_{x(k-1)\in\mathcal{X}} f(x(k)|x(k-1)) f(x(k-1)|z(1:k-1))$$

• Pamiętajmy, że niezależność warunkowa jest kluczowa dla istnienia rozwiązania.



Algorytm rekurencyjny cd.

• Uwzględnienie pomiaru (wykorzystując twierdzenie Bayesa) a posteriori

$$f(x(k)|z(1:k)) = f(x(k)|z(k), z(1:k-1))$$

$$= \frac{f(z(k)|x(k), z(1:k-1)) f(x(k)|z(1:k-1))}{f(z(k)|z(1:k-1))}$$

- z(k) oraz z(1/k-1) są warunkowo niezależne dla danego x(k)
- Normalizacja:

$$f(z(k)|z(1:k-1)) = \sum_{x(k)\in\mathcal{X}} f(z(k)|x(k)) f(x(k)|z(1:k-1))$$

• I wynik:
$$f(x(k)|z(1:k)) = \frac{f(z(k)|x(k)) \, f(x(k)|z(1:k-1))}{\displaystyle \sum_{\bar{x}(k) \in \mathcal{X}} f(z(k)|\bar{x}(k)) \, f(\bar{x}(k)|z(1:k-1))}$$



Algorytm rekurencyjny – podsumowanie

Krok 1: PREDYKCJA

$$f(x(k)|z(1:k-1)) = \sum_{x(k-1) \in \mathcal{X}} \overbrace{f(x(k)|x(k-1))}^{\text{model}} \overbrace{f(x(k-1)|z(1:k-1))}^{\text{poprzednia iteracja}}, \qquad k = 1, 2, \dots$$

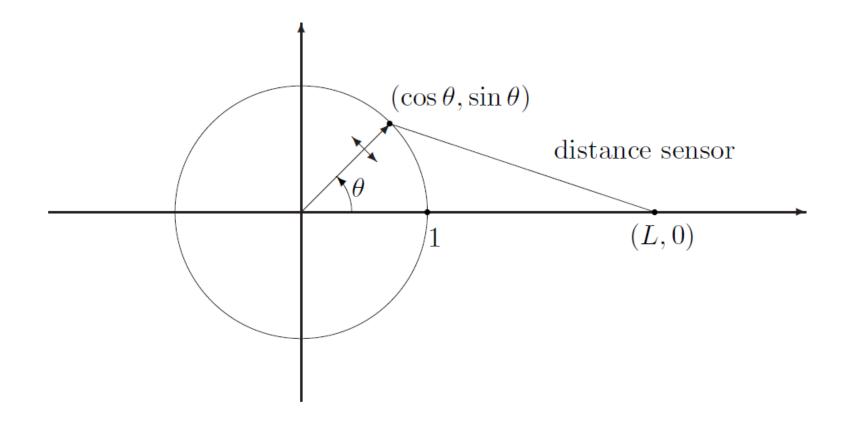
KROK 2: AKTUALIZACJA

$$f(x(k)|z(1:k)) = \underbrace{\frac{\int f(z(k)|x(k)) \int f(x(k)|z(1:k-1))}{\int f(z(k)|x(k)) \int f(x(k)|z(1:k-1))}}_{\bar{x}(k) \in \mathcal{X}} f(z(k)|\bar{x}(k)) f(\bar{x}(k)|z(1:k-1))$$



Algorytm rekurencyjny – przykład

- Obiekt porusza się losowo po okręgu
- Mierzymy odległość
- Chcemy poznać położenie



Algorytm rekurencyjny – przykład

- Obiekt porusza się w dyskretnych chwilach
 - Położenie

$$x(k) \in \{0, 1, \dots, N-1\}$$

I aktualny kąt:

$$\theta(k) = 2\pi \frac{x(k)}{N}$$

Dynamika dana jest przez:

$$x(k) = \operatorname{mod}(x(k-1) + v(k-1), N)$$

gdzie mod(.,N) oznacza modulo N operacji, tzn. mod(N,N) = 0 a mod(-1,N) = N-1.

Szum pomiarowy to:

$$v(k-1) = \begin{cases} 1 & \text{z prawdopodobieństwem } p \\ -1 & \text{z prawdopodobieństwem } 1-p \end{cases}$$



Algorytm rekurencyjny – przykład

• Model pomiaru odległości (w(k) szum pomiarowy w zakresie [-e, +e]):

$$z(k) = \sqrt{(L - \cos \theta(k))^2 + \sin^2 \theta(k)} + w(k)$$

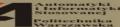
• Funkcja gęstości prawdopodobieństwa procesu i modelu pomiaru:

$$f(x(k)|x(k-1)) = \begin{cases} p & \text{if } x(k) = \text{mod}\big(x(k-1)+1, N\big) \\ 1-p & \text{if } x(k) = \text{mod}\big(x(k-1)-1, N\big) \\ 0 & \text{w pozostałych przypadkach} \end{cases}$$

$$f(z(k)|x(k)) = \begin{cases} \frac{1}{2e} & \text{if } \left| z(k) - \sqrt{(L - \cos \theta(k))^2 + \sin^2 \theta(k)} \right| \le e \\ 0 & \text{w pozostałych przypadkach} \end{cases}$$

• Inicjalizacja: f(x(0)) = 1/N







Filtr Kalmana jako obserwator stanu

Załóżmy, że obiekt jest stacjonarny:

$$x(k) = Ax(k-1) + u(k-1) + v(k-1)$$
 $x(0) \sim \mathcal{N}(x_0, P_0), v(k-1) \sim \mathcal{N}(0, Q)$
 $z(k) = Hx(k) + w(k)$ $w(k) \sim \mathcal{N}(0, R)$

Obserwowalność – para (A,H) jest obserwowalna jeśli macierz

$$\begin{bmatrix} H \\ HA \\ \vdots \\ HA^{n-1} \end{bmatrix}$$

• Jest rzędu *n*.



Filtr Kalmana jako obserwator stanu

Dla obiektów stacjonarnych filtr Kalmana jest zmienny, tzn.

$$P_{p}(k) = AP_{m}(k-1)A^{T} + Q$$

$$K(k) = P_{p}(k)H^{T}(HP_{p}(k)H^{T} + R)^{-1}$$

$$P_{m}(k) = (I - K(k)H)P_{p}(k).$$

- Co się dzieje z błędem estymacji $e(k) = x(k) \hat{x}_m(k)$ jak $k \to \infty$.
- Po uwzględnieniu równań i wariancji $P_p(k)$ otrzymujemy:

$$P_p(k+1) = AP_p(k)A^T + Q - AP_p(k)H^T (HP_p(k)H^T + R)^{-1}HP_p(k)A^T.$$



Filtr Kalmana jako obserwator stanu

- Obserwowalność jest wystarczająca do zbieżności
- Jeśli będziemy używać filtru Kalmana wystarczająco długo otrzymamy stacjonarny rozkład błędów estymacji



Wykrywalność

 System jest wykrywalny (detectable) wtedy i tylko wtedy, gdy wszystkie niestabilne mody są obserwowalne.

tzn.
$$\begin{bmatrix} A - \lambda I \\ H \end{bmatrix}$$
 jest pełnego rzędu dla wszystkich λ

- Idea wykrywalności:
 - Wykrywalność jest słabsza niż obserwowalność, tzn. (A,H) obserwowalne → (A,H) wykrywalne
 - Jeśli system jest wykrywalny, to wtedy:

$$\lim_{k\to\infty} P_P(k) = P_\infty$$



Jeśli wariancja filtru Kalmana jest zbieżna, to wtedy:

$$\lim_{k\to\infty} K(k) = K_{\infty}$$

- Zastosowanie stałej (niezmiennej w czasie) wartości K_{∞} zamiast zmiennej K(k)znacznie upraszcza implementację filtra.
- To się nazywa Filtr Kalmana w stanie ustalonym (Steady-State Kalman Filter)



- Wyznaczanie K_{∞}
 - Zakładając, że $P_P(k) \rightarrow P_{\infty}$

$$P_{\infty} = AP_{\infty}A^{T} + Q - AP_{\infty}H^{T}(HP_{\infty}H^{T} + R)^{-1}HP_{\infty}A^{T}.$$

- To jest równanie algebraiczne na P_{∞} , nazywane **Dyskretnym Algebraicznym Równaniem Riccatiego**
- Istnieją efektywne metody rozwiązania, w Matlabie dare (A', H', Q, R)
- Zatem wzmocnienie w stanie ustalonym dane jest wzorem:

$$K_{\infty} = P_{\infty}H^{T}(HP_{\infty}H^{T} + R)^{-1}$$



• Równanie Filtru Kalmana w stanie ustalonym z $\hat{x}(k) := \hat{x}_m(k)$:

$$\hat{x}(k) = (I - K_{\infty}H)A\,\hat{x}(k-1) + (I - K_{\infty}H)\,u(k-1) + K_{\infty}\,z(k)$$
$$= \hat{A}\,\hat{x}(k-1) + \hat{B}\,u(k-1) + K_{\infty}\,z(k),$$

liniowy stacjonarny system.

• Błąd estymacji:

$$e(k) = x(k) - \hat{x}(k) = \underbrace{(I - K_{\infty}H)A}_{\infty} e(k-1) + (I - K_{\infty}H) v(k-1) - K_{\infty} w(k)$$

!!!ważna stabilność!!!

bo inaczej nie zbiega



- Co może się nie udać:
 - $P_P(k)$ nie zbiega jak $k \to \infty$,
 - $P_P(k)$ zbiega, ale do innego rozwiązania dla różnych $P_P(1)$. Nie wiadomo co przyjąć do wyliczeń K_∞ ,
 - $(I K_{\infty}H)A$ jest niestabilne.

•





Filtr Kalmana w stanie ustalonym - uwagi

- Filtr Kalmana jest optymalnym estymatorem (system liniowy i rozkłady Gaussowskie) bez względu na obserwowalność i wykrywalność.
- Obserwowalność i wykrywalność są cechą systemu a nie algorytmu estymacji.
- Obserwowalność i wykrywalność można również zdefiniować dla systemów niestacjonarnych czy też nieliniowych. Ale wtedy warunki są odmienne ...



Rozważmy nieliniowy system dyskretny w czasie:

$$x(k) = q_{k-1}(x(k-1), u(k-1), v(k-1))$$
 $E[x(0)] = x_0, Var[x(0)] = P_0$ $E[v(k-1)] = 0, Var[v(k-1)] = Q(k-1)$ $z(k) = h_k(x(k), w(k))$ $E[w(k)] = 0, Var[w(k)] = R(k)$

- q_{k-1} jest różniczkowalna względem x(k-1) i v(k-1)
- h_k jest różniczkowalna względem x(k) i w(k)





- O co chodzi z różniczkowaniem EKF?
- Aby otrzymać estymatę stanu nieliniowego systemu linearyzujemy równania stanu i stosujemy standardową predykcję filtru Kalmana a potem aktualizację do zlinearyzowanych równań.



- Załóżmy, że wyznaczono $\hat{x}_m(k-1)$ oraz $P_m(k-1)$ jako aproksymację średniej oraz wariancji stanu x(k-1) na podstawie pomiarów x(1:k-1).
- Linearyzacja równania stanu wokół punktu $x(k-1)=\hat{x}_m(k-1)$ oraz

$$v(k-1) = E[v(k-1)] = 0$$
 prowadzi do:

$$\begin{aligned} x(k) &\approx q_{k-1}(\hat{x}_m(k-1), u(k-1), 0) \\ &+ \underbrace{\frac{\partial q_{k-1}(\hat{x}_m(k-1), u(k-1), 0)}{\partial x} \cdot (x(k-1) - \hat{x}_m(k-1)) + \underbrace{\frac{\partial q_{k-1}(\hat{x}_m(k-1), u(k-1), 0)}{\partial v} \cdot v(k-1)}_{=:L(k-1)} \\ &= A(k-1)x(k-1) + \underbrace{L(k-1)v(k-1)}_{=:\tilde{v}(k-1)} + \underbrace{q_{k-1}(\hat{x}_m(k-1), u(k-1), 0) - A(k-1)\hat{x}_m(k-1)}_{=:\tilde{v}(k-1)} \\ &= A(k-1)x(k-1) + \tilde{v}(k-1) + \xi(k-1), \end{aligned}$$
 szum procesowy znane wejście



• Szum procesowy $\tilde{v}(k-1)$ ma zerową wartość średnią oraz wariancję:

$$Var[\tilde{v}(k-1)] = L(k-1)Q(k-1)L^{T}(k-1)$$

I otrzymujemy

$$\begin{split} \hat{x}_p(k) &= A(k-1)\hat{x}_m(k-1) + \xi(k-1) \\ &= q_{k-1}(\hat{x}_m(k-1), u(k-1), 0) \qquad \text{wstawiając } \xi(k-1) \\ P_p(k) &= A(k-1)P_m(k-1)A^T(k-1) + L(k-1)Q(k-1)L^T(k-1). \end{split}$$



Linearyzacja prowadzi do

$$\begin{split} z(k) &\approx h_k(\hat{x}_p(k),0) + \underbrace{\frac{\partial h_k(\hat{x}_p(k),0)}{\partial x}}_{=:H(k)} \cdot (x(k) - \hat{x}_p(k)) + \underbrace{\frac{\partial h_k(\hat{x}_p(k),0)}{\partial w}}_{=:M(k)} \cdot w(k) \\ &= H(k)x(k) + \underbrace{M(k)w(k)}_{=:\tilde{w}(k)} + \underbrace{h_k(\hat{x}_p(k),0) - H(k)\hat{x}_p(k)}_{=:\zeta(k)} \\ &= H(k)x(k) + \tilde{w}(k) + \zeta(k), \end{split}$$

• $\widetilde{w}(k)$ ma zerową wartość średnią oraz wariancję:

$$Var[\tilde{v}(k-1)] = M(k)R(k)M^{T}(k)$$





Zastosowanie aktualizacji filtru Kalmana do równania zlinearyzowanego prowadzi do:

$$K(k) = P_p(k)H^T(k) \left(H(k)P_p(k)H^T(k) + M(k)R(k)M^T(k) \right)^{-1}$$

$$\hat{x}_m(k) = \hat{x}_p(k) + K(k) \left(z(k) - H(k)\hat{x}_p(k) - \zeta(k) \right)$$

$$= \hat{x}_p(k) + K(k) \left(z(k) - h_k(\hat{x}_p(k), 0) \right) \quad \text{wstawiając } \xi(k-1)$$

$$P_m(k) = \left(I - K(k)H(k) \right) P_p(k).$$



- 1. Inicjalizacja $\hat{x}_m(0) = x_0, P_m(0) = P_0$
- 2. Krok 1 predykcja *a priori*

$$\hat{x}_p(k) = q_{k-1}(\hat{x}_m(k-1), u(k-1), 0)$$

$$P_p(k) = A(k-1)P_m(k-1)A^T(k-1) + L(k-1)Q(k-1)L^T(k-1)$$

gdzie

$$A(k-1) := \frac{\partial q_{k-1}(\hat{x}_m(k-1), u(k-1), 0)}{\partial x} \quad i \qquad L(k-1) := \frac{\partial q_{k-1}(\hat{x}_m(k-1), u(k-1), 0)}{\partial v}.$$



1. Aktualizacja *a posteriori*

$$K(k) = P_p(k)H^T(k) \left(H(k)P_p(k)H^T(k) + M(k)R(k)M^T(k) \right)^{-1}$$

$$\hat{x}_m(k) = \hat{x}_p(k) + K(k) \left(z(k) - h_k(\hat{x}_p(k), 0) \right)$$

$$P_m(k) = \left(I - K(k)H(k) \right) P_p(k)$$

gdzie

$$H(k) := \frac{\partial h_k(\hat{x}_p(k), 0)}{\partial x}$$
 i $M(k) := \frac{\partial h_k(\hat{x}_p(k), 0)}{\partial w}$.



Uwagi:

- Macierze A(k-1), L(k-1), H(k) i M(k) otrzymujemy linearyzując równania stanu wokół punktu aktualnego (!zależy od pomiarów!). Zatem EKF nie można wyznaczyć off-line.
- Aktualny stan i szum są bliskie aktualnie mierzonym a szumu mają wartość oczekiwaną zero. Wtedy linearyzacja jest dobrą aproksymacją nieliniowej wartości.
 - Ale to może być błędne założenie! Bądź co bądź Gaussowski szum jest nieograniczony.
- Aproksymaty stanu i wariancji są tylko aproksymatami, jako że proces nieliniowy jest. I nie muszą wystarczająco odzwierciedlać rzeczywistych wartości, szczególnie gdy nieliniowości są znaczące. Niemniej stosujemy, jako że lepszych nie mamy!
- Pomimo tego, że EKF jest tylko bardzo zgrubna Bayesowską estymatą stanu i nie można zagwarantować zbieżności, EKF często stosowany jest praktycznie.



Hybrydowy EKF

- A co zrobić w przypadku, gdy proces jest ciągły w czasie a pomiary są zbierane w dyskretnych chwilach czasu.
- Proces:

$$\dot{x}(t) = q(x(t), u(t), v(t), t)$$

$$z[k] = h_k(x[k], w[k]) \qquad \qquad \text{E}[w[k]] = 0, \text{Var}[w[k]] = R,$$

Oznaczamy dyskretyzację:

$$x[k] := x(kT)$$
, T – okres próbkowania.

- Istnieją dwa rozwiązania:
 - 1. Dyskretyzujemy proces i dalej postępujemy jak wcześniej.
 - 2. Podejście bezpośrednie.





Hybrydowy EKF

- Uwagi:
 - Wymagane jest różniczkowanie równań różniczkowych stanu (np. metodami Runge-Kutte). Jest to kosztowne obliczeniowo. Podsumowując różniczkowanie numeryczne równań to podstawowy koszt dodatkowy.
 - Analogiczne kwestie wystąpią dla standardowe filtru Kalmana proces liniowy.
 - Podejście można rozszerzyć dla liniowego modelu zarówno procesu jak i pomiarów.

