copyright: M. Dwornik
dwornik@agh.edu.pl

Splot, korelacja i PACF

1 Splot

Splotem funkcji f(t) i g(t) nazywamy funkcję h(t):

$$h(t) = f(t) * g(t) = \int_{-\infty}^{\infty} f(\tau) \cdot g(t-\tau) d\tau \tag{1}$$

$$h[n] = f[n] * g[n] = \sum_{m=-\infty}^{\infty} f[m] \cdot g[n-m] \tag{2}$$

Funkcja ta ma następujące własności:

- jest przemienna:

$$f(t) * g(t) = g(t) * f(t) \Leftrightarrow \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau = \int_{-\infty}^{\infty} g(\tau)f(t-\tau)d\tau \tag{3}$$

- jest łączna:

$$f(t) * \Big(g(t) * h(t)\Big) = \Big(f(t) * g(t)\Big) * h(t) \tag{4}$$

- Delta Diraca jest elementem neutralnym dla splotu:

$$\delta(t) * f(t) = f(t) \tag{5}$$

- jest rozdzielcza względem dodawania:

$$f(t) * \Big(g(t) + h(t)\Big) = f(t) * g(t) + f(t) * h(t) \tag{6}$$

2 Implementacja splotu

W pakiecie MathWorks MatLAB splot wektorów realizowany jest poprzez funkcję
```
z = conv(x, y, 'shape');
```
gdzie parametr `'shape'` przyjmuje następujące wartości:

- `full` - zwraca całą długość wektora splotu (wartość domyślna):
  `length(z) = length(x)+length(y)-1;`
- `same` - wektor splotu ma rozmiar pierwszego argumentu:
  `length(z) = length(x);`
- `valid` - zwraca tylko tą część wektora splotu, która została policzona bez wykorzystania wartości zerowych krótszego sygnału dodanych na brzegach:
  `length(z) = max(length(x)-max(0,length(y)-1),0);`

3 Filtracja liniowa

Filtracja liniowa w domenie czasu jest to splot sygnału wraz z maską filtra. Rozmiar maski na ogół jest liczbą nieparzystą, gdzie wartość sumomnożenia przypisywana jest elementowi centralnemu. W zależności od sumy elementów maski wyróżniamy:

- filtrację dolnoprzepustową (uśredniającą), gdzie suma elementów maski równa się **1** oraz wszystkie elementy maski są nieujemne, np.:
  LP1=[0.2 0.2 0.2 0.2 0.2]; LP2=[0.25 0.5 0.25]; sygnał Gaussa;
- filtrację górnoprzepustową, gdzie suma elementów maski równa się **0**, np:
  HP1=[-1 0 1]; czy HP2=[-1 2 -1];

4 Filtracja nieliniowa

Filtracja nieliniowa, czyli w najprostszym ujęciu, taka, której nie można wyrazić za pomocą splotu. Przykładem takiej filtracji jest filtracja medianowa - `medfilt1(x, N)`, gdzie każdemu punktowi centralnemu przypisywana jest mediana z wartości pokrytych maską.

4.1 Filtracja adaptatywna (Wienera)

Osobną kategorią filtrów nieliniowych są filtry adaptacyjne. Są to filtry, których charakterystyka zmienia się w zależności od analizowanego obszaru. W celu usuwania szumu działają one dwuetapowo:

1. Dla każdego punktu i jego otoczenia obliczamy wartość parametru, który kwalifikuje dany punkt jako należący lub nie do krawędzi. W przypadku filtru Wienera jest to wariancja w obszarze obrazu pokrytym maską.
2. jeżeli dany punkt został zakwalifikowany jako nie należący do krawędzi, zostaje on poddany silnemu uśrednieniu. W przeciwnym wypadku jego wartość pozostaje bez zmian lub poddany zostaje uśrednieniu o niewielkiej mocy.

Do filtrów tego typu zaliczamy filtr Wienera: `wiener2(obraz, [rozmiar])`.

4.2 Filtr Savitzky-Golay

Metoda Savitzky-Golay polega na lokalnym wygładzaniu sygnału przy użyciu wielomianów określonego stopnia. Do filtracji służy polecenie `sgolayfilt(sygnał, stopień, rozmiar)`, gdzie rozmiar musi być nieparzystą, dodatnią liczbą całkowitą.

5 Korelacja

Korelacja dwóch funkcji f(t) i g(t) dana jest następującym wzorem:

$$R_{fg}(t) = \int_{-\infty}^{\infty} f(\tau) \cdot g^*(\tau - t) d\tau \tag{7}$$

$$R_{fg}[n] = \sum_{k=-\infty}^{\infty} f[k] \cdot g^*[k-n] \tag{8}$$

gdzie g^* - sprzężenie liczby zespolonej. Korelacja sygnału z samym sobą nosi nazwę autokorelacji:

$$R_{ff}(t) = \int_{-\infty}^{\infty} f(\tau) \cdot f^*(\tau - t) d\tau \tag{9}$$

Do najważniejszych własności funkcji korelacji zaliczamy:

- Symetryczność:

$$R_{fg}(t) = R^*_{gf}(-t), \quad R_{ff}(t) = R^*_{ff}(-t) \tag{10}$$

- dla dowolnego przesunięcia τ wartość bezwzględna funkcji autokorelacji jest mniejsza lub równa funkcji autokorelacji dla zerowego przesunięcia:

$$|R_{ff}(\tau)| \leqslant R_{ff}(0) \tag{11}$$

- Energia sygnału f(t) równa jest jego autokorelacji dla zerowego przesunięcia:

$$E_f = R_{ff}(0) = \int_{-\infty}^{\infty} |f(t)|^2 dt \tag{12}$$

- $$\int_{-\infty}^{\infty} R_{fg}(\tau) d\tau = \left(\int_{-\infty}^{\infty} f(t)dt\right)\left(\int_{-\infty}^{\infty} g^*(t)dt\right) \tag{13}$$

6 Implementacja korelacji

W pakiecie MatLAB korelacja jest realizowana poprzez polecenie:
`[z, przes] = xcorr(x, y, okno, 'normalizacja');`
Autokorelacja dana jest poleceniem:
`[z, przes] = xcorr(x, 'normalizacja');`.
W obu przypadkach zmienna `przes` jest opcjonalna i zawiera wartości przesunięcia dla funkcji korelacji. Dodatkowy parametr `'normalizacja'` opisuje, czy oraz jak wartości współczynników korelacji ma znormalizowane (`biased`, `unbiased`, `none` - domyślne, `coeff` - tylko dla autokorelacji, współczynniki normalizowane są względem zerowego przesunięcia). Parametr `okno` decyduje o wielkości okna korelacji i jednocześnie determinuje rozmiar wektora wyjściowego (`N*okno + 1`). W przypadku braku tego parametru przyjmowana jest wartość `max(length(x), length(y))`, a krótszy wektor uzupełniany jest zerami do rozmiaru dłuższego wektora.

7 Funkcja autokorelacji cząstkowej (PACF)

Jest to autokorelacja z przesunięciem (opóźnieniem), gdzie nie uwzględnia się wpływu autokorelacji pochodzącej od krótszych przesunięć. Jest to metoda stosowana do identyfikacji modelu. Model autoregresywny (AR - ang. *autoregressive*) służy głównie do modelowania szeregów, których widmo jest prążkowe (np. suma składowych harmonicznych). Model ten można zapisać w postaci:

$$x_t = \varphi_0 + \sum_{k=1}^{p} \varphi_k \cdot x_{t-k} + \epsilon_t \tag{14}$$

gdzie: φ_0 - stała, φ - parametry modelu, ϵ - biały szum (o średniej 0 i wariancji σ^2).

Jeżeli szereg x_t jest stacjonarny o postaci z równania 14 to spełnia on twierdzenie Yule-Walkera dla $\forall n \in \mathrm{N} \cup 0$:

$$\rho_n = \sum_{k=1}^{p} \rho_k \cdot \rho_{n-k} \tag{15}$$

gdzie ρ_n - funkcja autokorelacji. Wartości te można policzyć korzystając z funkcji : `ar = autocorr(x)`.
Wykorzystując powyższą zależność oraz, że $\rho(0) = 1$, estymacja parametrów sprowadza się do rozwiązania układu równań liniowych:

$$\begin{bmatrix} 1 & \rho(1) & \cdots & \rho(p-1) \\ \rho(1) & 1 & \cdots & \rho(p-2) \\ \rho(2) & \rho(1) & \cdots & \rho(p-3) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(p-1) & \rho(p-2) & \cdots & 1 \end{bmatrix} \cdot \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ \vdots \\ \varphi_p \end{bmatrix} = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \rho(3) \\ \vdots \\ \rho(p) \end{bmatrix} \Rightarrow \mathbf{R} \cdot \varphi = \mathbf{r} \tag{16}$$

Macierz $\mathbf{R}$ jest symetryczna, dodatnio zdefiniowana, wszystkie jej wartości własne są dodatnie i nieosobliwa.
Przybliżając wartość autokorelacji funkcjami kowariancji:

$$\rho_k \approx \frac{r_k}{r_o} = \frac{\mathrm{cov}(x_t, x_{t+k})}{\mathrm{cov}(x_t, x_t)} = \frac{\sum_t \left((x_t - \mu) \cdot (x_{t+k} - \mu)\right)}{\sum_t \left((x_t - \mu)^2\right)} \tag{17}$$

gdzie μ to wartość oczekiwana szeregu x_t, dostajemy wzór na wartości parametrów modelu:

$$\begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ \vdots \\ \varphi_p \end{bmatrix} = \left(\begin{bmatrix} 1 & r_1 & \cdots & r_{p-1} \\ r_1 & 1 & \cdots & r_{p-2} \\ r_2 & r_1 & \cdots & r_{p-3} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p-1} & r_{p-2} & \cdots & 1 \end{bmatrix} \right)^{-1} \cdot \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \vdots \\ r_p \end{bmatrix} \Rightarrow \varphi = \mathbf{R^{-1}} \cdot \mathbf{r} \tag{18}$$

Macierz $\mathbf{R}$ jest symetryczna; dodatnio zdefiniowana, ma nieujemne wartości własne i istnieje do niej macierz odwrotna.

Wariancję σ^2 estymujemy korzystając ze wzoru:

$$\sigma^2 = \mathrm{cov}(0) \cdot \left(1 - \sum_{k=1}^{p} r_k \cdot \theta_k\right) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2 \left(1 - \sum_{k=1}^{p} r_k \cdot \theta_k\right) \tag{19}$$

Estymacja wartości funkcji PACF wymaga ponownie użycia wzorów Yule-Walkera.

$$\phi_{hh} = \frac{\det \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{h-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{h-3} & \rho_2 \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{h-4} & \rho_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{h-1} & \rho_{h-2} & \rho_{h-3} & \cdots & \rho_1 & \rho_h \end{bmatrix}}{\det \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{h-2} & \rho_{h-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{h-3} & \rho_{h-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{h-4} & \rho_{h-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{h-1} & \rho_{h-2} & \rho_{h-3} & \cdots & \rho_1 & 1 \end{bmatrix}} \tag{20}$$

I tak dla wartości funkcji PACF dla modelu AR wynoszą:

- dla AR(1):

$$\phi_{11} = \frac{\rho_1}{1} \tag{21}$$

- dla AR(2):

$$\phi_{22} = \frac{\det \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{bmatrix}}{\det \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix}} \tag{22}$$

copyright: M. Dwornik
dwornik@agh.edu.pl

- dla AR(3):

$$\phi_{33} = \frac{\det \begin{bmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{bmatrix}}{\det \begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix}} \tag{23}$$

Wyniki można porównać korzystając z funkcji w MatLAB:
`pacf = parcorr(x, Method = "yule-walker");`

8 Zadania

1. Policz w pakiecie MatLAB splot dwóch funkcji prostokątnych ($|t| \leqslant 5,\ Fs = 100$):

$$f(t) = g(t) = \begin{cases} 1 & |t| \leqslant 2 \\ 0 & |t| > 2 \end{cases} \tag{24}$$

2. Policz (**ręcznie**) splot dwóch wektorów:
   `A = [-1, 0, 2,- 4]` oraz `B = [2, 0, 1]`.
   `C = [1 2 2 3 3 3]` oraz `D = [-1 0 1]`.
   `E = [0 0 1 0 0]` oraz `F = [0.25 0.5 0.25]`.
   `G = [1+i 1-i 0 3i]` oraz `H = [2i 1 1+i]`.

3. Policz splot (**ręcznie oraz w programie MatLAB**) funkcji prostokątnej $f(t)$ o parametrach: amplituda=1.0, szerokość=3.0, środek=0.5; z wektorem `[-1, 0, 1]` dla t = -10 : 1 : 10.
   Wyświetl wykresy funkcji oraz wynik ich splotu.

4. Policz (**w programie MatLAB**) splot funkcji gaussowskiej f(t) o parametrach (sr=0.0, std=0.2, Amp=1) z funkcją g(t) dla $n \in \mathbb{N}$) dla $t \in \langle -5, 15 \rangle$ spróbkowanych z $F_S = 50Hz$. Ćwiczenie powtórz dla różnych wartości odchylenia standardowego funkcji f(t).

$$g(t) = \text{Ш}(t) = \begin{cases} 1 & dla \quad t \in \mathbb{N} \\ 0 & dla \quad t \notin \mathbb{N} \end{cases} \tag{25}$$

5. Policz (**w programie MatLAB**) splot funkcji harmonicznej $f(t)$ o amplitudzie $A_1 = 1.0$, okresie $\omega_1 = 0.5s$ i czasie trwania $T_1 = 10s$ z funkcją prostokątną g(t) o czasie trwania $T_2 = 0.2s$ i amplitudzie $A_2 = 1/(T_2 \cdot F_S)$. Oba sygnały spróbkowane są z częstotliwością $F_S = 100Hz$. Stwórz wykresy sygnałów oraz ich splotu. Zbadaj wpływ dlugości T_2 na wynik splotu.

6. Wygeneruj poziomy sygnał dyskretny o $t \in \langle 0, 5 \rangle$s i Fs=200Hz, który będzie wypełniony liczbami z rozkładu normalnego N(0, 0.5). Zaimplementuj splot z wektorem `ones(1,N)/N`, dla N będącego liczbą nieparzystą. Stwórz wykres zmienności odchylenia standardowego sygnału wynikowego w zależności od wartości N.

7. Policz (**ręcznie**) korelację i konwolucję wektorów:
   `A=[ 1 -1.5 -2 1 ]` i `B=[ 2 -2 1 ]`.

copyright: M. Dwornik
dwornik@agh.edu.pl

`C=[1, 2, -3, 1]` i `D=[1, 0, -1]`.
`E=[1+i, 2, -3i, -1]` i `F=[1-i, 2i, 3]`.
Wynik sprawdź w pakiecie MatLAB.

8. Policz korelację dwóch sygnałów trójkątnych o amplitudzie 1.0:
f(t): środek = 2, szerokość = 2 oraz
g(t): środek = 0, szerokość = 3 dla $t \in < -5, 5 >$ z $F_S = 100Hz$.

9. Stwórz sygnał będący sumą trzech sygnałów ($t \in \langle 0, 10\rangle,\ Fs = 100$) typu sinc:
- amplituda A = 2 dla t = 2, okres = 1;
- amplituda A = 1 dla t = 3.5, okres = 1;
- amplituda A = 0.5 dla t = 4.5, okres = 1.
Dla takiego sygnału policz i wyświetl wartości funkcji autokorelacji. Obliczenia powtórz używając zamiast funkcji sinc, funkcji harmonicznych o stałej amplitudzie.

10. Wczytaj sygnał z pliku *corr_ 01.txt*. Przy użyciu funkcji korelacji znajdź początek trwania sygnału:

   - prostokątnego o amplitudzie 0.7 i czasie trwania 7s;
   - sygnału trójkątnego o amplitudzie 0.8 i czasie trwania 10s;
   - piłokształtnego o amplitudzie 1.0 i czasie trwania 5s;
   - sygnału gaussowskiego o amplitudzie 0.8, std=$2\sqrt{2}$, i czasie trwania 20s;

   Na podstawie pliku (1 kolumna - czas [s], 2 kolumna - amplituda sygnału) określ częstotliwość próbkowania.

11. Wczytaj sygnał z pliku *corr_ 02.txt*. Przy użyciu funkcji korelacji znajdź początek trwania sygnału:

   - prostokątnego o amplitudzie 0.8 i czasie trwania 10s;
   - obu sygnałów trójkątnych o amplitudzie 1.0 i czasie trwania 10s;
   - każdego z ząbków funkcji piłokształtnej o amplitudzie 1.0 i czasie trwania 5s;
   - środka (wierzchołka) funkcji gaussowskiego o amplitudzie 0.8, std=5.

12. Policz wartości funkcji autokorelacji i PACF dla następujących wektorów:

   - x = [1, 2, -3, 2, -2] dla modelu AR(2);
   - x = [-3, 1, 0, 3 ,-1] dla modelu AR(2);
   - x = [2, 1, 2, 1 ,-1] dla modelu AR(2) i AR(3);
   - x = [-2, 1, 3, -3 ,1] dla modelu AR(3).

   Wyniki sprawdź w MatLAB.