



MI

Metody Identyfikacji

wykład #1

1. *Wstęp*

Kwestie formalne

- Wykład (55 pkt.):
 - sprawdzian zaliczeniowy (2 godz.) – 25 maja 2020
 - sprawdzian poprawkowy (2 godz.) – 8 czerwca 2020
- Projekt (nieograniczona ilość punktów):
 - Zadanie: dane przemysłowe
 - obrona (konieczna): 1 czerwca 2020
 - Ocena za projekt składa się z trzech składowych:
 - Prowadzący wykład: 0-15 pkt.
 - Prowadzący projekt: 0-20 pkt.
 - Każdy Zespół studencki ma do rozdysponowania po 10 punktów pomiędzy inne Zespoły

Kwestie formalne

- Warunki zaliczenia przedmiotu:
 - **wykład:** ponad 27 pkt.
 - **projekt:** dostarczenie w terminie (1 czerwca 2020, godz. 16:15) sprawozdania
- Skala ocen:
 - <50.5 = 2
 - 50.5-60 = 3
 - 60.5-70 = 3.5
 - 70.5-80 = 4
 - 80.5-90 = 4.5
 - 90.5-100 = 5

Literatura

- Zdzisław Bubnicki: *Identyfikacja obiektów sterowania*, PWN, 1974.
- Michael S. Mahoney: Historical Perspectives on Models and Modeling
- Lenart Ljung: *System Identification: Theory for Users*
- Chris Bissell and Chris Dillon (Eds.): *Ways of Thinking, Ways of Seeing Mathematical and Other Modelling in Engineering and Technology*, Springer, 2012.
- Rolf Isermann and Marco Münchhof: *Identification of Dynamic Systems*

Pośrednie wprowadzenie

1. Nate Silver: *The signal and the noise. The art and science of prediction*
2. Nassim Nicholas Taleb: *Statistical Consequences of Fat Tails: Real World Preasymptotics, Epistemology, and Applications*

pośrednie wprowadzenie

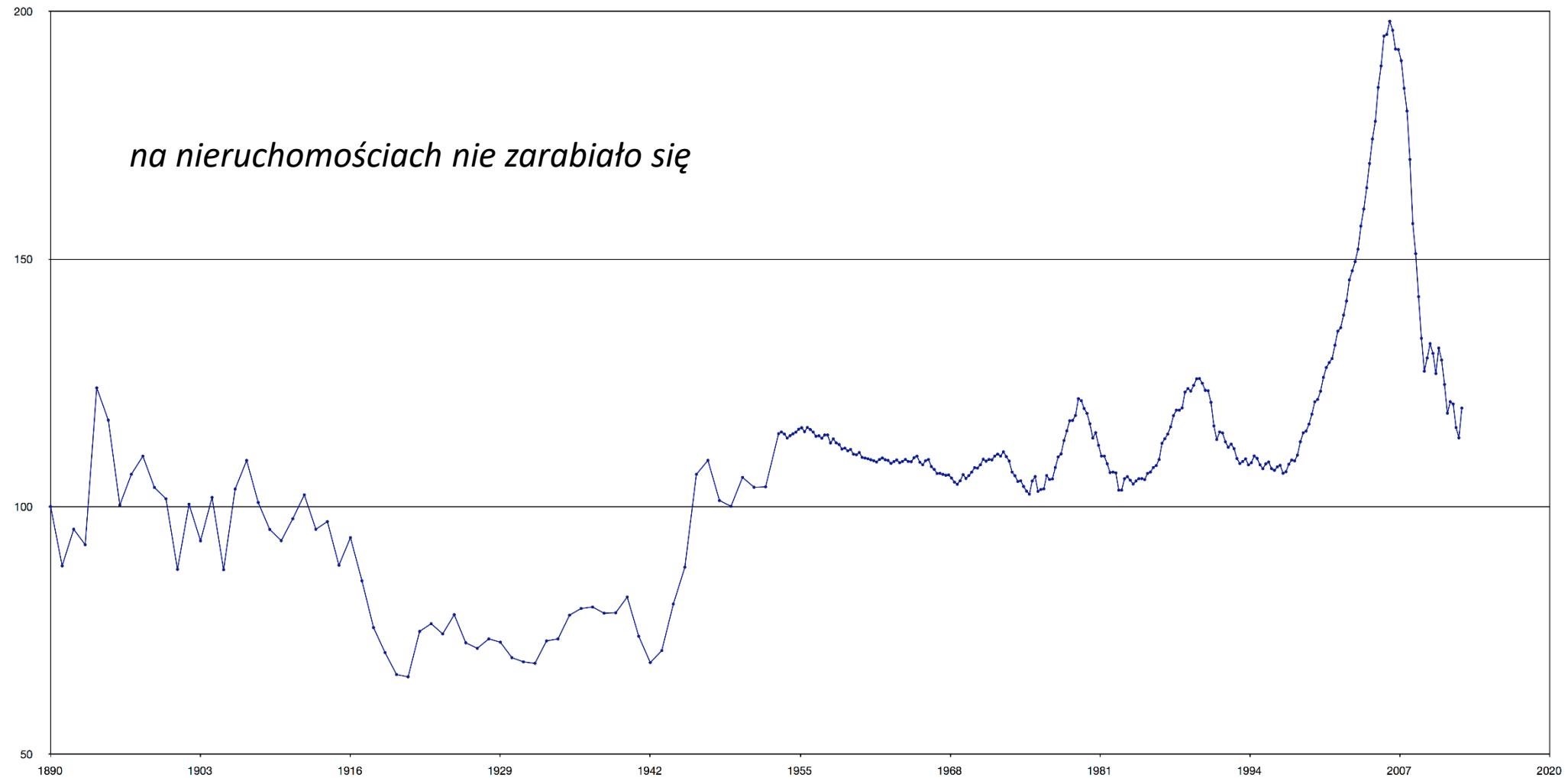
Nate Silver: *The signal and the noise. The art and science of prediction.*

- A Catastrophic Failure of Prediction

We focus on those signals that tell a story about world as we would like it to be, not how it really is. We ignore the risks that are hardest to measure, even when they pose the greatest threats to our well-being.

- Bańka na rynku nieruchomości
- Agencje ratingowe dawały najwyższą ocenę AAA dla tych inwestycji
 - Zależność od siebie poszczególnych kredytów hipotecznych – poszczególne narzędzia i zdarzenia nie były od siebie niezależne
 - Mnóstwo powiązanych instrumentów pochodnych
 - Wszyscy wiedzieli, nikt nic nie zrobił – nikomu nie zależało na jakości prognoz – agencje zarabiały na ilości prognoz
- Różnica pomiędzy *ryzykiem* a *niepewnością*.
 - **Ryzyko:** coś co możesz wyrazić za pomocą jakieś ceny (Frank Knight, 1921)
 - **Niepewność:** niemierzalne

Akt 1: ceny nieruchomości



Ciąg dalszy

- Akt 2: dźwignia
 - Zwykły Amerykanin miał 65% dobrobytu związanego z nieruchomością, natomiast oszczędności małe
 - Dla każdego **1\$** zainwestowanego w hipoteki istniało co najmniej **50\$** w instrumentach pochodnych na giełdzie
 - dodatnie sprzężenie zwrotne
 - podaż-popyt → sprzężenie ujemne, stabilizuje rynek
 - strach-pożądanie → sprzężenie dodatnie, destabilizacja
- Akt 3: Bezrobocie
 - Kryzys finansowy generuje bezrobocie

Year	Without stimulus (%)	With stimulus (%)	Actual (%)
2009	7.0	7.0	7.0
2010	9.5	8.5	10.1
2011	8.5	7.5	9.0
2012	7.0	6.0	8.2
2013	6.7	5.7	-

Sources: Bureau of Labor Statistics; White House

Sytuacja jakiej nie było wcześniej: *out of sample*

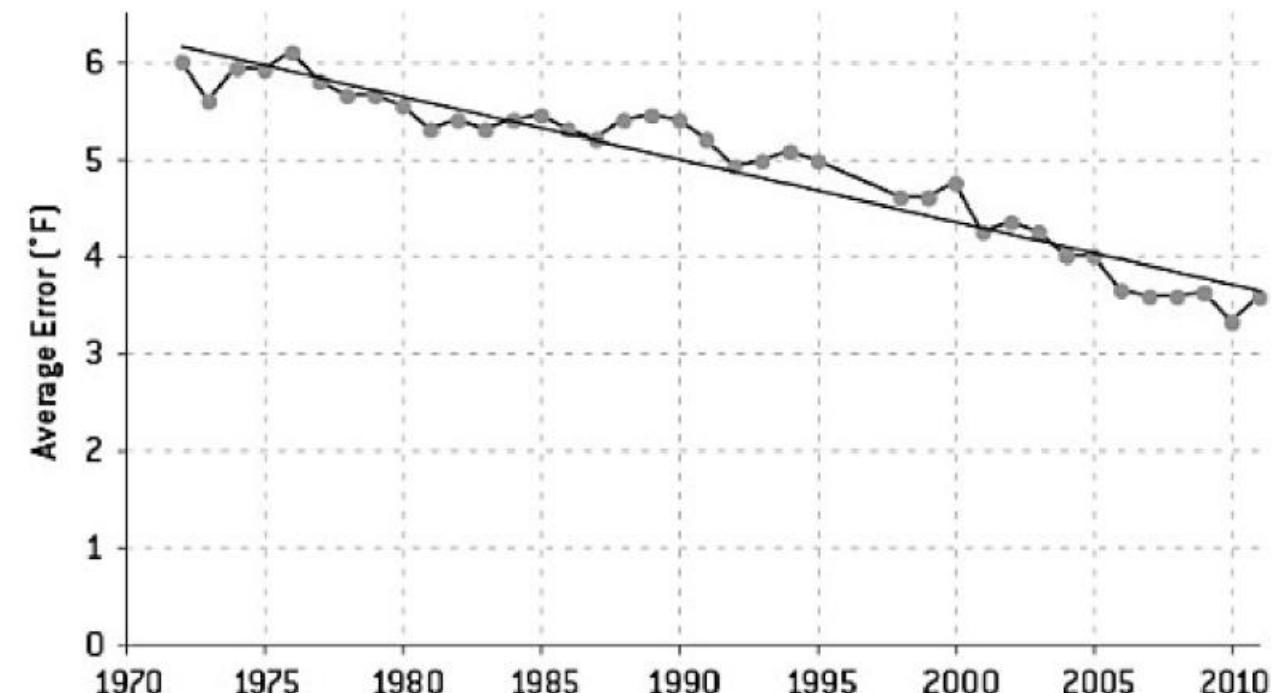
Model jest tylko przybliżeniem rzeczywistości – musimy znać jago zakres stosowalności

Are you Smarter than a Television Pundit?

- *Mądrze głowy w telewizji zawsze mylą się w swoich prognozach*
 - 15% zdarzeń, które nie miały szans na zdarzenie, wystąpiło
 - 25% zdarzeń, które były absolutnie pewne, nigdy nie wydarzyło się
- Hedgehogs and foxes
The fox knows many little things, but the hedgehog knows one big thing. Archilochus
- Błąd w myśleniu: *dobra prognoza nie ulega zmianie*
- Wykorzystujmy każdą nadarzającą się informację, również tą jakościową – podejście hybrydowe

For Years You've Been Telling us that Rain is Green

- Prognozy pogody są bardzo dobre, wraz za zwiększającą się mocą obliczeniową następuje poprawa prognoz – podejście podziału dziedziny w 3D
- Ekspert modyfikuje modele i prognozy – lepiej widzimy wzorce w szumie
- Kalibracja
- Właściwości dobrej prognozy:
 - Dokładność
 - Rzetelność
 - Wartość ekonomiczna



Desperately Seeking Signal

- W przeciwieństwie do pogody nie jesteśmy w stanie przewidywać trzęsień ziemi
 - nie ma świętego Grala
- Różnica pomiędzy prognozą a przewidywaniem
- *Power Law* – prawo potęgowe
- Poszukiwanie sygnału w szumie
- Overfitting
- Ekstrapolacja

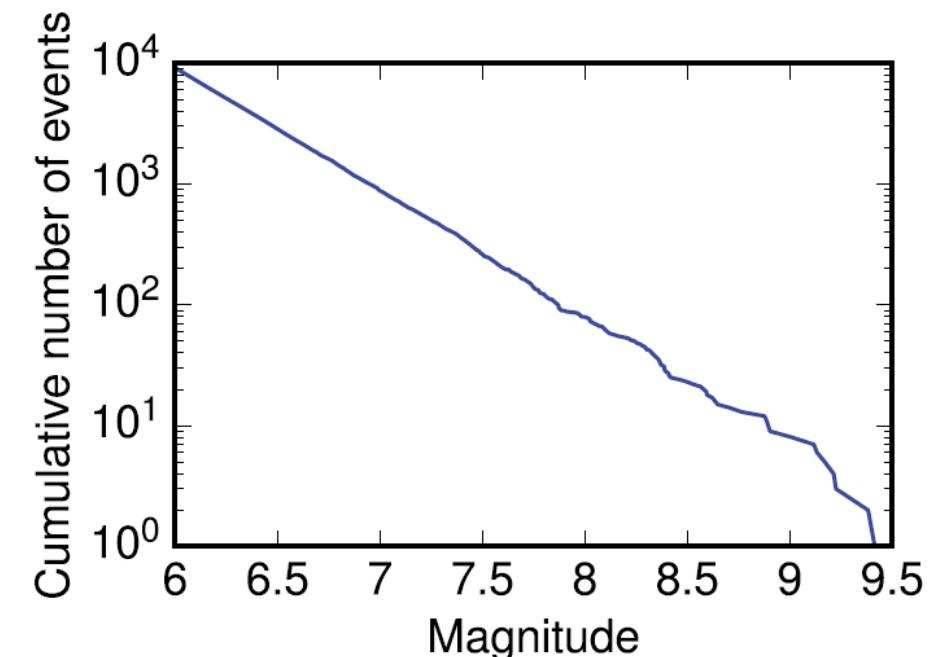
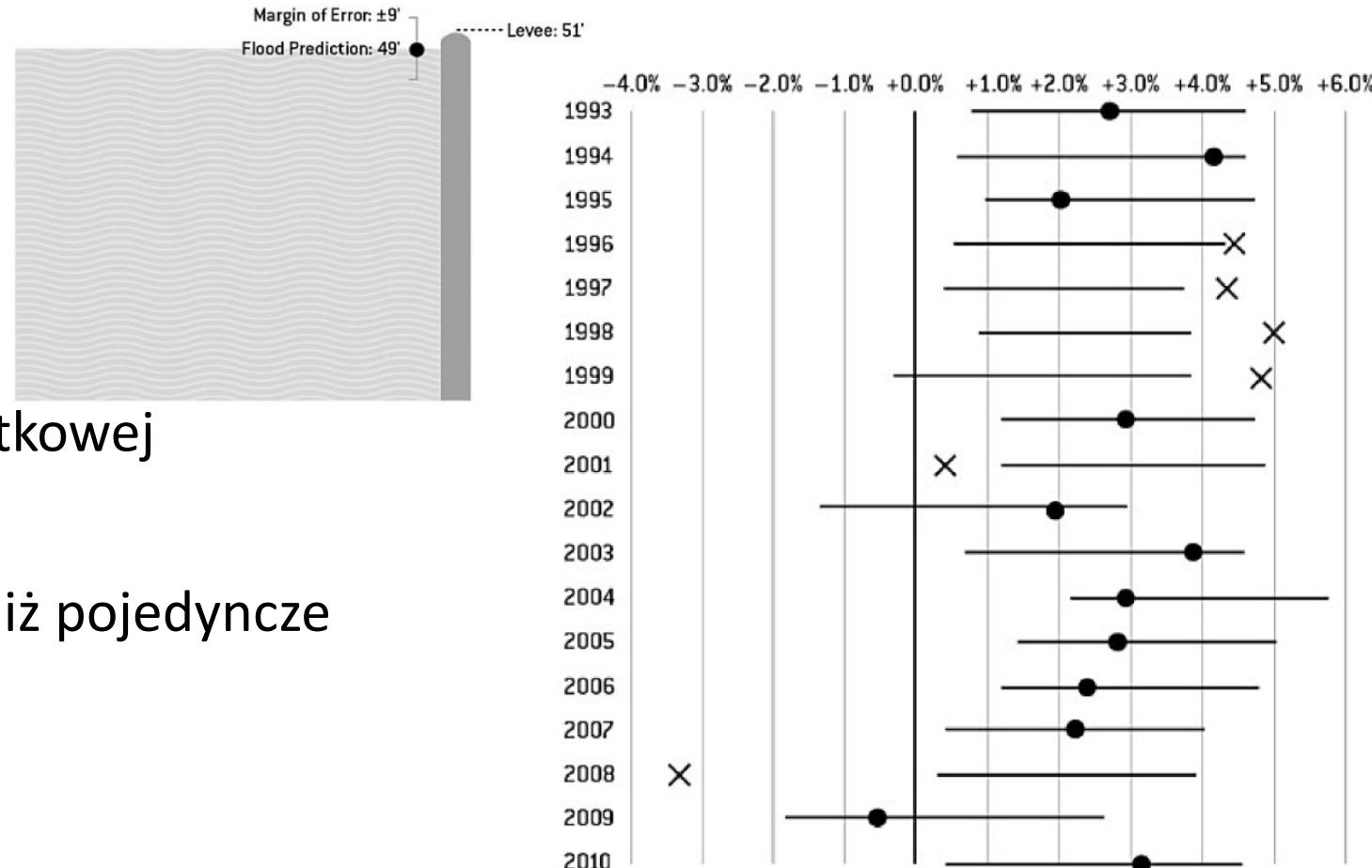


Figure 3. Magnitude-frequency distribution used in the synthetic data sets. Events follow a Gutenberg-Richter distribution with $b = 1$ and a minimum magnitude of $M_{\min} = 6$. While this differs from the empirical value of $b = 1.26$ observed for the PAGER/PDE catalog, using $b = 1$ makes determining the number of events at a given magnitude level more straightforward. Alternative magnitude-frequency distributions with $b = 0.8$, $b = 1.2$, and $M_{\min} = 5$ are also considered for the ETAS simulations.

How to Drown in Three Feet of Water

- Przekleństwo statystyki
- Prognozy PKB – dokładność
- Mnóstwo danych
- Mnóstwo korelacji
- Brak zależności przyczynowo - skutkowej
- Jak powinna być struktura modelu
- Prognozy zagregowane są lepsze niż pojedyncze



Role models

- Prognozowanie rozprzestrzeniania chorób / pandemii
- Niebezpieczeństwo ekstrapolacji
- Prognozy samospełniające się
 - Głosowanie na kandydata wygrywającego
 - Prognozowanie modnego koloru na przyszły rok
 - Choroby
- Prognozy samo-kasujące się
 - Wszyscy używają jednego systemu nawigacji w celu ominięcia korka
- Paradoks HIV w San Francisco 1990-2000
- Modele agentowe
 - *All models are wrong, but some models are useful*
 - *The best model of a cat is a cat.*

Less and Less and Less Wrong

to err and err and err again, but less and less and less, Piet Hein

- prawdopodobieństwem *a priori*

- Dotyczy zdarzenia, które jeszcze nie nastąpiło, a my pytamy o prawdopodobieństwo zajścia jakiegoś zdarzenia przed jego wykonaniem, czyli *a priori*.
- Liczymy to prawdopodobieństwo zgodnie ze wzorem na prawdopodobieństwo całkowite
- Mówiąc krótko, znamy wszystkie możliwe przyczyny, a nie znamy skutku (rezultatu).
- [Ronald Aymer Fisher \(próbkowanie\)](#)

- Prawdopodobieństwem *a posteriori*

- W tym przypadku doświadczenie już zostało wykonane, znamy jego wynik. Nas interesuje prawdopodobieństwo tego, że dana przesłanka przyczyniła się do danego skutku.
- To jest sytuacja *a posteriori*, znamy skutek, a pytamy o prawdopodobieństwo, że to właśnie ta określona przyczyna spowodowała skutek.
- [Thomas Bayes](#)

pośrednie wprowadzenie

Nassim Nicholas Taleb: *Statistical Consequences of Fat Tails*

- *The main idea behind the Incerto project is that while there is a lot of uncertainty and opacity about the world, and an incompleteness of information and understanding, there is little, if any, uncertainty about what actions should be taken based on such an incompleteness, in any given situation.*
- *Complication without insight: the clarity of mind of many professionals using statistics and data science without an understanding of the core concepts, what it is fundamentally about.*

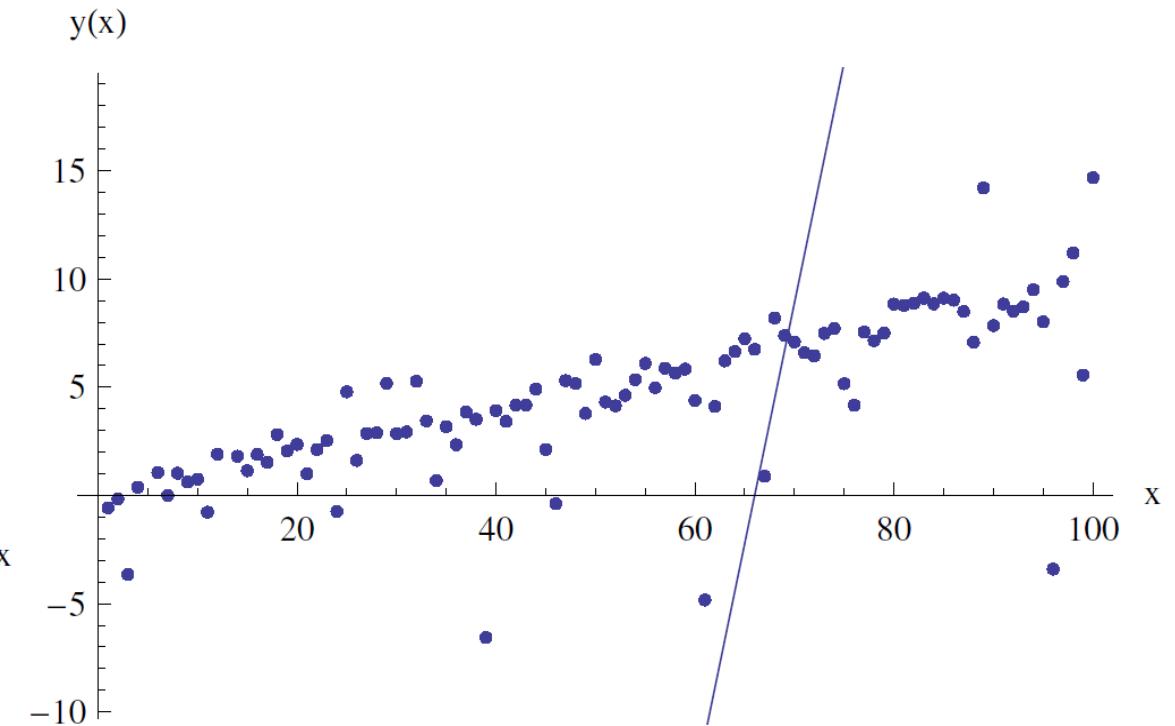
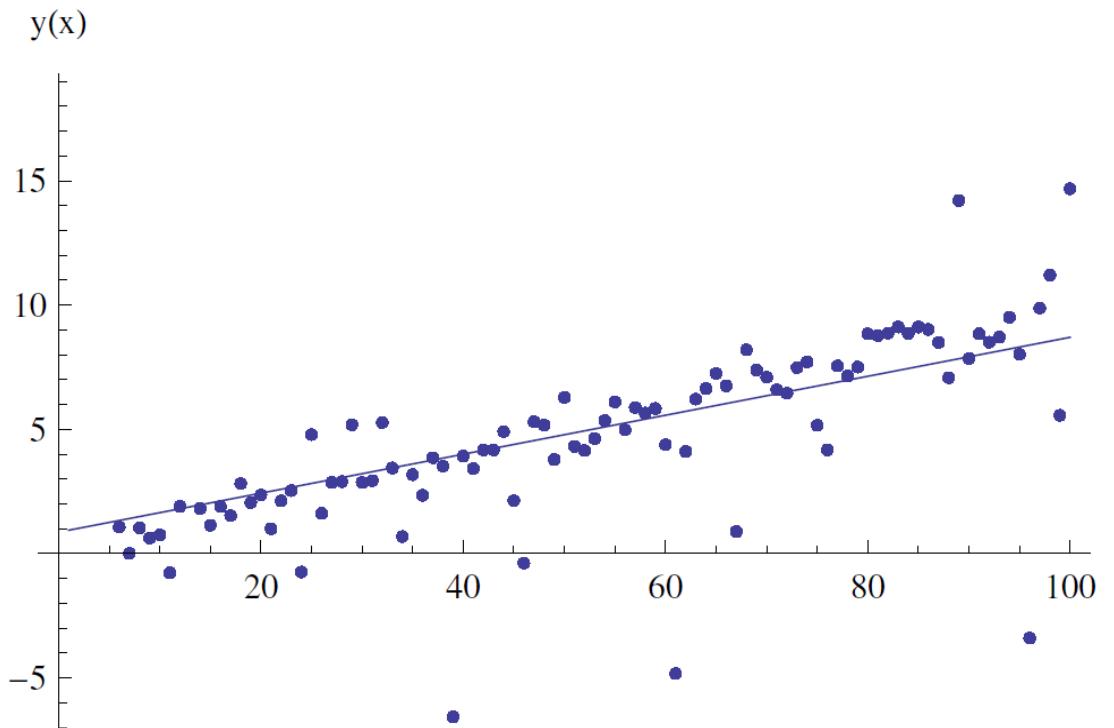
Pinker problem

- Discussion of "evidence" that is not statistically significant, or use of metrics that are uninformative because they do not apply to the random variables under consideration – like for instance making inferences from the means and correlations for fat tailed variables. This is the result of:
 - the focus in statistical education on Gaussian or thin-tailed variables,
 - the absence of probabilistic knowledge combined with memorization of statistical terms,
 - complete cluelessness about dimensionality.
- Example of pseudo-empiricism: comparing death from terrorist actions or epidemics such as ebola (fat tailed) to falls from ladders (thin tailed).

Cienkie a grube ogony

- In Mediocristan, when a sample under consideration gets large, no single observation can really modify the statistical properties.
 - the probability of sampling higher than X twice in a row is greater than sampling higher than $2X$ once
- In Extremistan, the tails (the rare events) play a disproportionately large role in determining the properties.
 - the probability of sampling higher than $2X$ once is greater than the probability of sampling higher than X twice in a row.

regresja



obserwacje

1. *The law of large numbers, when it works, works too slowly in the real world.*
2. *The mean of the distribution will rarely correspond to the sample mean; it will have a persistent small sample effect (downward or upward) particularly when the distribution is skewed (or one-tailed).*
3. *Metrics such as standard deviation and variance are not useable.*
4. *Practically every single economic variable and financial security is thick tailed. Of the 40,000 securities examined, not one appeared to be thin-tailed. This is the main source of failure in finance and economics.*
5. *Practically any paper in economics using covariance matrices is suspicious.*
6. *Linear least-square regression doesn't work (failure of the Gauss-Markov theorem).*
7. *Maximum likelihood methods can work well for some parameters of the distribution (good news).*

obserwacje, cd.

8. *The gap between disconfirmatory and confirmatory empiricism is wider than in situations covered by common statistics i.e., the difference between absence of evidence and evidence of absence becomes larger.*
9. *The method of moments (MoM) fails to work. Higher moments are uninformative or do not exist.*
10. *There is no such thing as a typical large deviation*

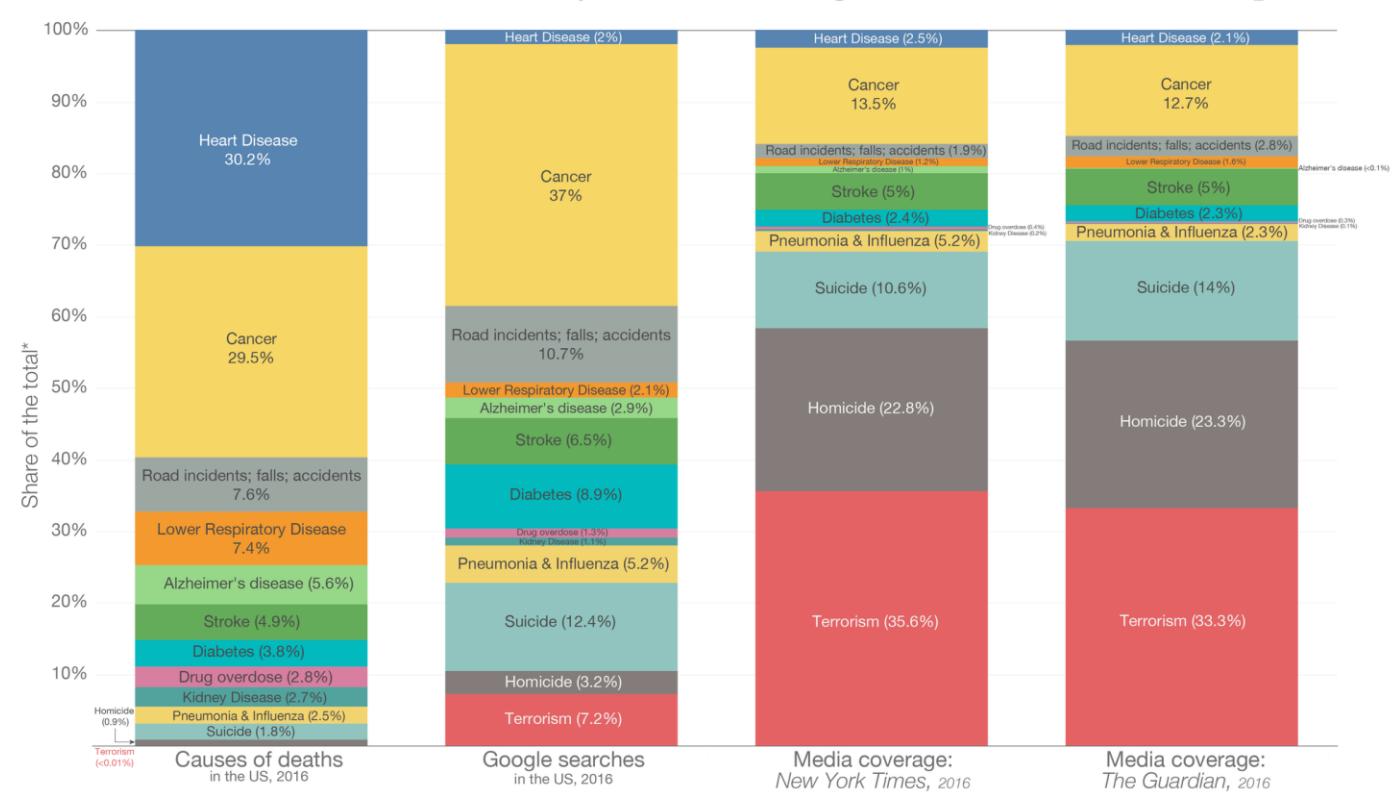
Epistemologia

- We do not observe probability distributions, just realizations.
- A probability distribution cannot tell you if the realization belongs to it.
- You need a meta-probability distribution to discuss tail events (i.e., the conditional probability for the variable to belong to a certain distributions vs. others).

potoczność

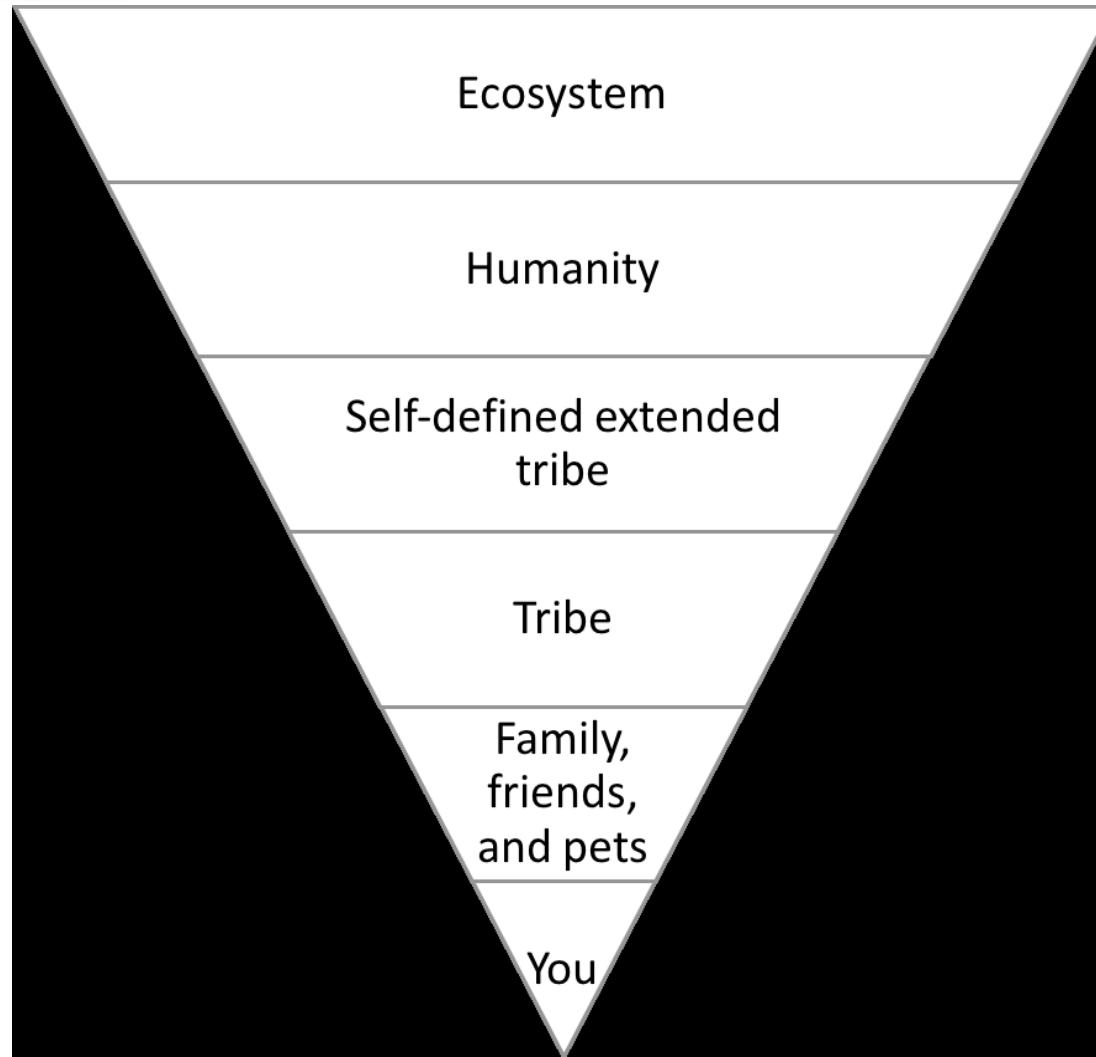
Causes of death in the US

What Americans die from, what they search on Google, and what the media reports on



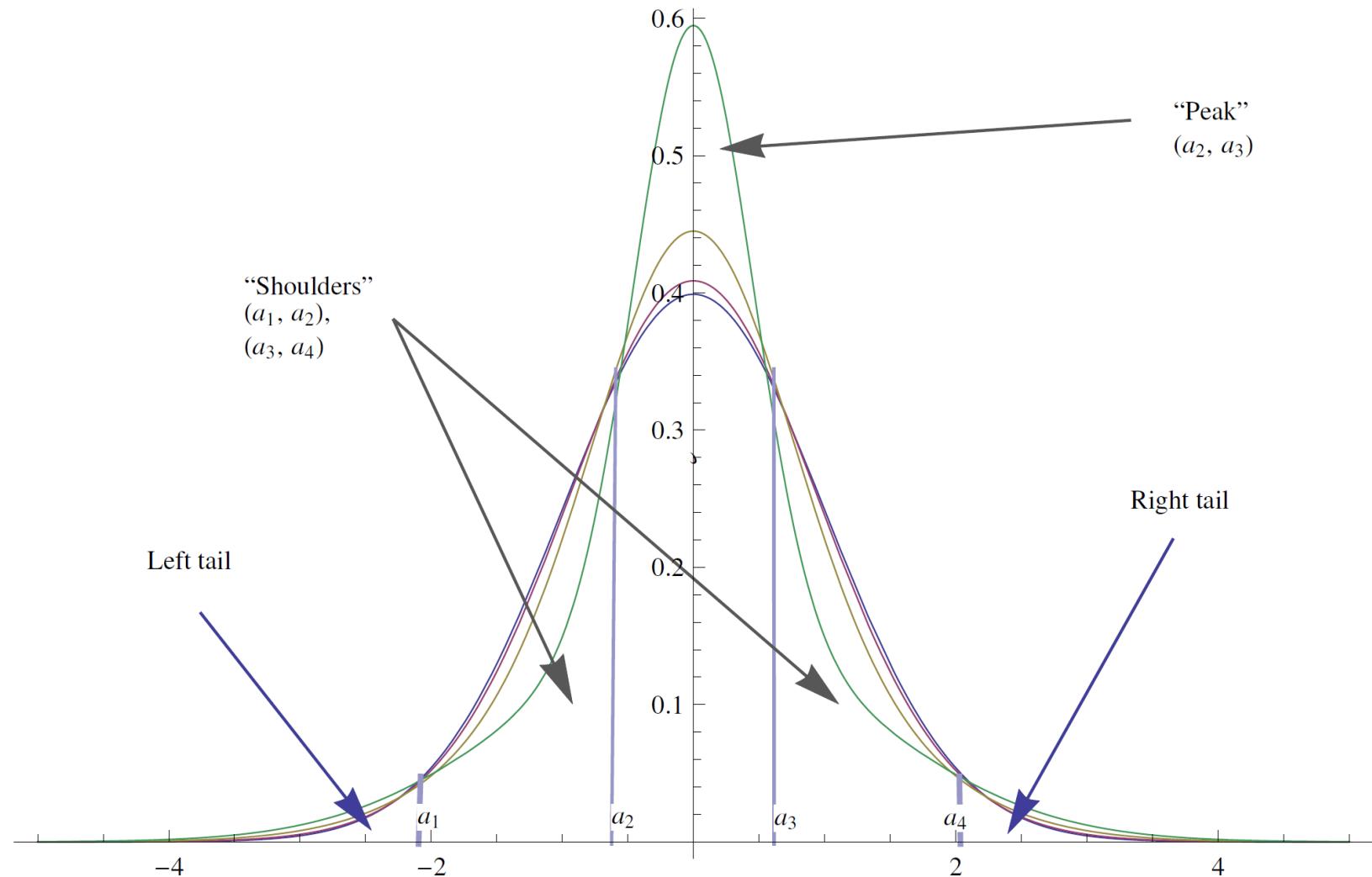
Bill Gates's Naive (Non-Statistical) Empiricism: the founder of Microsoft is promoting and financing the development of the above graph, yet at the same time claiming that the climate is causing an existential risk, not realizing that his arguments conflict since existential risks are necessarily absent in past data. Furthermore, a closer reading of the graphs shows that cancer, heart disease, and Alzheimer, being ailments of age, do not require the attention on the part of young adults and middle-aged people something terrorism and epidemics warrant.

Another logical flaw is that terrorism is precisely low because of the attention it commands. Relax your vigilance and it may go out of control. The same applies to homicide: fears lead to safety.

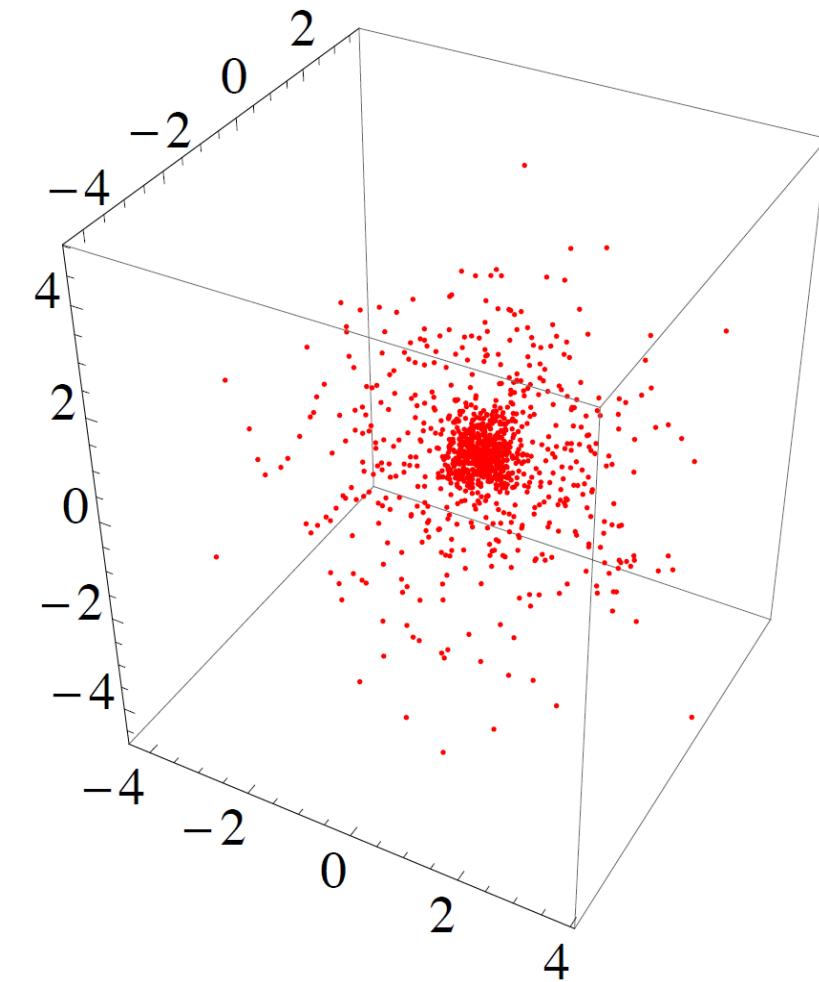
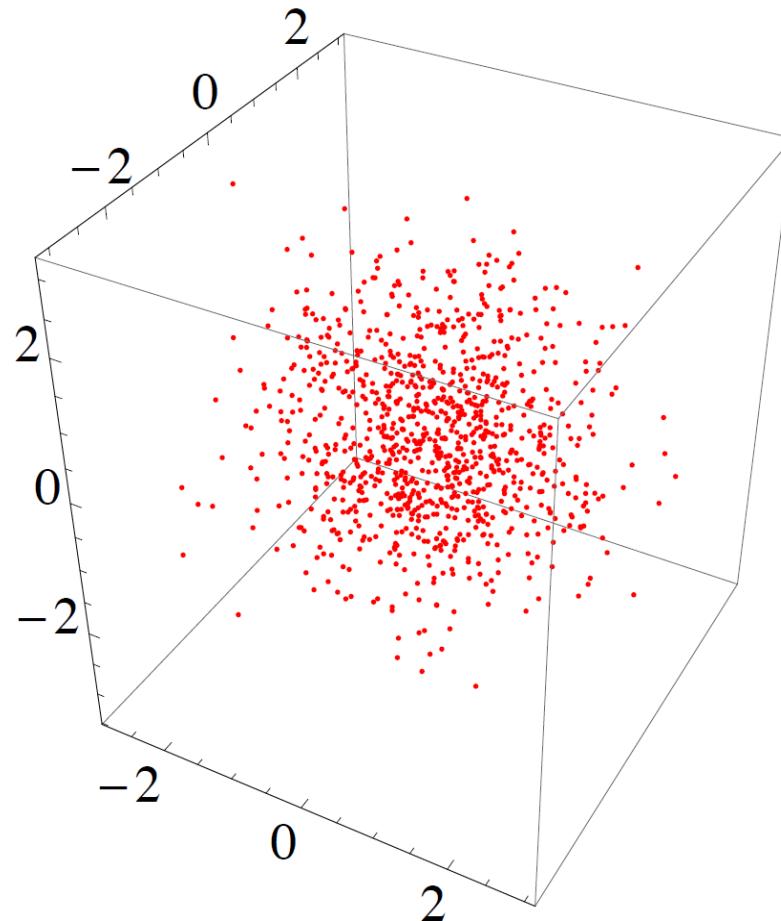


*A hierarchy for survival.
Higher entities have a longer
life expectancy, hence tail risk
matters more for these.
Lower entities such as you
and I are renewable.*

Gdzie zaczyna się ogon?



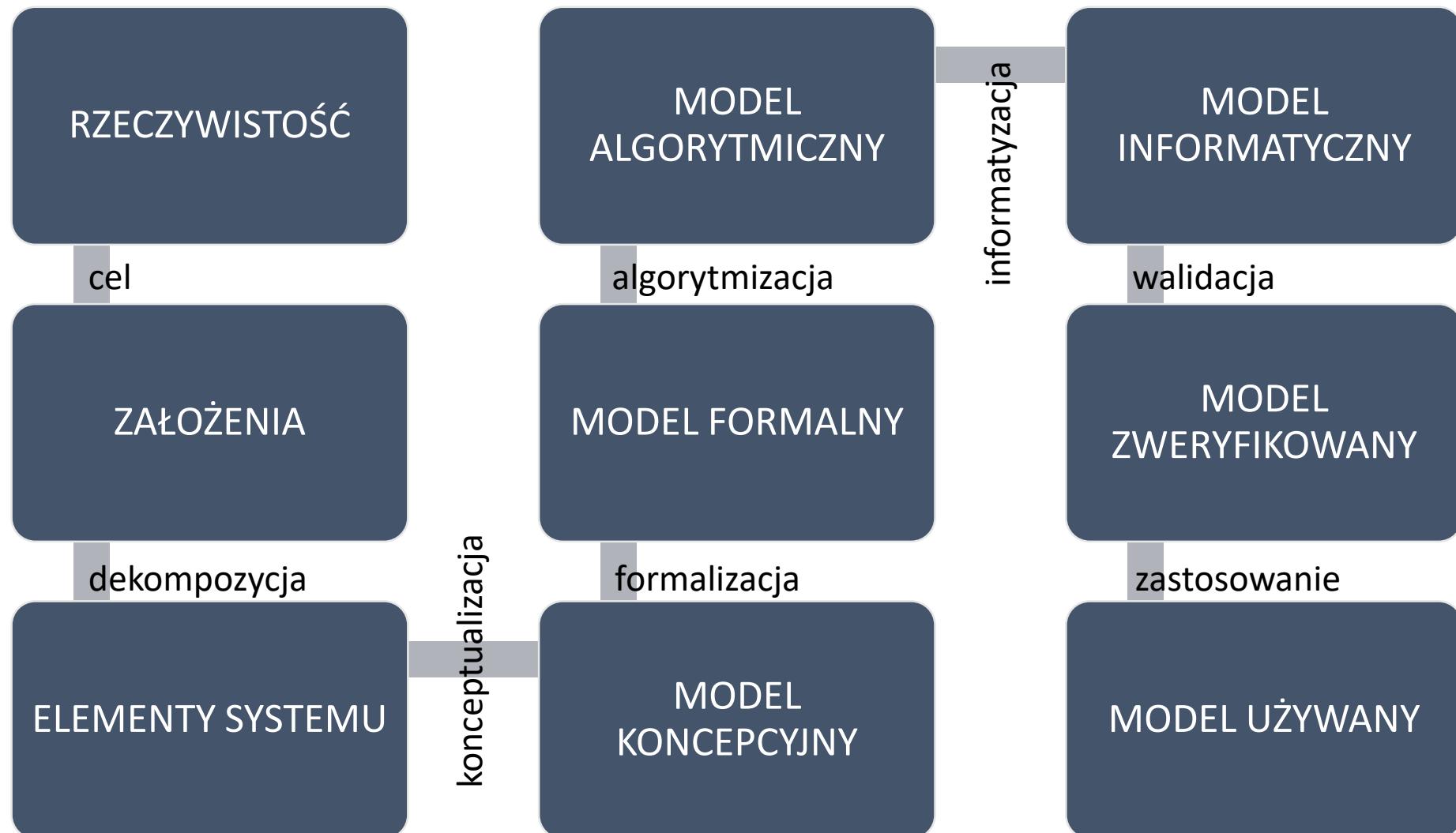
Ogony a wymiary



Inne kwestie

- Jak wiele danych potrzeba do analizy?
- Wartości ekstremalne, outliers, ogony

Proces modelowania



Identyfikacja

- **IDENTYFIKACJA:** wygenerowanie z istniejącego fragmentu rzeczywistości elementów niezbędnych do zbudowania modelu z punktu widzenia potrzeby, możliwości, warunków i ograniczeń itp.
- **PYTANIA:**
 - Jakie elementy tworzą system?
 - Jakie relacje tworzą strukturę?
 - Jakie elementy mi relacje są istotne ze względu na cel?
 - Jak struktura systemu wpływa na funkcję systemu?
 - Jak otoczenie może wpływać na zmianę struktury?
- **PROBLEMY:**
 - Funkcje i procesy realizowane w systemie
 - Zachowanie się systemu w danych warunkach
 - Organizacja realizacji procesów w systemie
 - Uzyskanie pożądanego przebiegu procesów
 - Związki cech opisujących system z badaną właściwością

Identyfikacja

- KONCEPTUALIZACJA: transformacja zbioru założeń otrzymanych w procesie identyfikacji do postaci modelu określającego relacje między zidentyfikowanymi elementami i ustalającymi atrybuty niezbędne do opisu systemu.
- FORMALIZACJA: budowa modelu formalnego (matematycznego).
- ALGORYTMIZACJA: przekształcenie modelu matematycznego w postać numeryczną (algorytm, schemat blokowy).
- INFORMATYZACJA: budowa modelu komputerowego (programu komputerowego).
- WERYFIKACJA: konfrontacja oszacowań otrzymanych w modelu z oszacowaniami otrzymanymi w warunkach naturalnych (rzeczywistych).
- ADAPTACJA: ustalenie zakresu, warunków zastosowania, możliwości posługiwania się modelem (opracowanie instrukcji posługiwania się modelem).

Identyfikacja

- Identyfikacja zajmuje się wyznaczaniem modeli matematycznych obiektów.
 - Model matematyczny definiuje w sposób ścisły zachowanie się obiektu w określonych warunkach.
 - Warunki te są określone przez wejścia i wyjścia obiektu w chwili obecnej i w przeszłości.
 - Model jest pewnym przybliżeniem rzeczywistego obiektu, idealizacją z ograniczającymi założeniami (ograniczona ilość wejść i wyjść, liniowość).
 - Stosuje się różne postacie modeli w zależności od przeznaczenia tworzonego modelu i struktury identyfikowanego obiektu.
 - Identyfikacja jest przeprowadzana na podstawie informacji pomiarowej o wielkościach wejściowych i wyjściowych obiektu.
 - W odniesieniu do sygnałów również stosuje się często idealizację, przybliżając rzeczywiste sygnały ich idealnymi odpowiednikami opisywanymi matematycznie.
 - Wstępny etapem identyfikacji jest określenie charakteru obiektu na podstawie rejestracji jego sygnałów (statyczny czy dynamiczny, jeśli dynamiczny to jakiego rzędu, o jakiej dynamice).

Klasyfikacja modeli

- Liniowe ↔ Nieliniowe
 - Niezbyt dokładne (jakościowe) ↔ Precyzyjne (ilościowe)
 - Przyczynowo - skutkowe ↔ Pozostałe
 - Statyczne ↔ Dynamiczne
 - Deterministyczne ↔ Stochastyczne
 - O stałych skupionych ↔ Rozproszone (o stałych rozłożonych)
 - Stacjonarne ↔ Niestacjonarne
 - z czasem dyskretnym ↔ z czasem ciągłym
 - Fizykochemiczne ↔ Empiryczne
 - SISO ↔ MIMO
- analiza wymiarowa ↔

Identyfikacja (jako proces)

- Nieinwazyjna (nie zaburza) ↔ • Inwazyjna
- strukturalna ↔ • parametryczna
- ilościowa ↔ • jakościowa
- statyczna ↔ • dynamiczna
- jednorazowa ↔ • powtarzalna, okresowa, rekurencyjna
- *a priori* ↔ • *a posteriori*

Identyfikacja

- Procedura identyfikacji jest procedurą iteracyjną:
 - Planowanie pomiarów w celu uzyskania danych. Często uzyskanie danych wystarczających do identyfikacji jest możliwe tylko na drodze odpowiedniego pobudzania wejść układu, a takie oddziaływanie nie zawsze jest możliwe w praktyce.
 - Pozyskanie i weryfikacja danych obiektowych
 - Wybór struktury modelu, tzn. określenie postaci i rzędu równań, opóźnień itp.
→ Identyfikacja strukturalna
 - Estymację parametrów modelu. Metoda estymacji zależy od przyjętego modelu i tego, czy obliczenia są prowadzone w trybie on- czy off-line.
→ Identyfikacja parametryczna
 - Sprawdzenie (walidacja) poprawności modelu. Do sprawdzenia najlepiej wykorzystać dane pomiarowe inne niż wykorzystane do estymacji parametrów. W razie negatywnego wyniku testu procedurę powtarza się.

Sztuka prognozowania

- **Przewidywanie** - wnioskowanie o zdarzeniach nie znanych na podstawie zdarzeń znanych – **co się może wydarzyć**:
 - racjonalne (logiczny proces)
 - zdroworozsądkowe (wnioskowanie oparte na doświadczeniu),
 - naukowe (wnioskowanie z wykorzystaniem reguł nauki),
 - nieracjonalne (wróżby i proroctwa).
- **Prognozowanie (predykcja)** - racjonalne, naukowe przewidywanie przyszłych zdarzeń (zazwyczaj wnioskowanie oparte na danych czasowych lub przekrojowych) – **co się wydarzy i kiedy**.
 - sformułowany z wykorzystaniem dorobku nauki
 - odnoszący się do określonej przyszłości
 - weryfikowalny empirycznie w przyszłości
 - wartość logiczna sądu nieznana w momencie jego formułowania
 - niepewny, ale akceptowany
 - sąd oznajmujący, a nie warunkowy
- Co prognozujemy/przewidujemy?

Cele prognozowania

- Preparacyjna
- Aktywizująca
- Informacyjna

Klasyfikacja prognoz

- Perspektywa czasowa
 - Prognozy krótkookresowe
 - Prognozy średniookresowe
 - Prognozy długookresowe
- Cel
 - Prognoza badawcza (w tym ostrzegawcza)
 - Prognoza realistyczna
- Skutek
 - Prognoza samorealizująca się
 - Prognoza samounicestwiająca się

Błąd prognozy

- Ex ante
 - Błąd modelu na danych historycznych, np. błąd średniokwadratowy
- Ex post
 - Trafność prognozy (jej jakość, skuteczność, powtarzalność, stosowalność)

Zagrożenia

- chętnych do dzielenia się swoimi wizjami nie brakuje - dostajemy to czego chcieliśmy i stanowczo zbyt wiele opieramy na tych prognozach
- skłonność do mylenia pewności siebie z biegłością w temacie
- zwykle nie jesteśmy w stanie ocenić wartości prognoz
- tworzenie prognoz tak nieprecyzyjnych, rozmytych i nieokreślonych w czasie, że wręcz niemożliwych do zweryfikowania
- nikt nie jest nieomylny
- wartości odstające - *outliers*



MI

Metody Identyfikacji

wykład #2

1. *Wprowadzenie*
 1. *Sygnały stochastyczne*
 2. *Sygnały okresowe*

Sygnały okresowe

- Klasyczny sygnał sinusoidalny

$$x(t) = x_0 \sin(\omega_0 t + \alpha) \text{ with } \omega_0 = \frac{2\pi}{T_0}$$

- Funkcja autokorelacji (wystarczy całkować na połowie okresu)
 - ACF sinusa jest funkcją cosinus

$$R_{xx}(\tau) = \frac{2x_0^2}{T_0} \int_0^{\frac{T_0}{2}} \sin(\omega_0 t + \alpha) \sin(\omega_0(t + \tau) + \alpha) dt = \frac{x_0^2}{2} \cos \omega_0 \tau$$

- I otrzymujemy gęstość widmową

$$\begin{aligned} S_{xx}(\omega) &= \frac{x_0^2}{2} \int_{-\infty}^{\infty} \cos \omega_0 \tau \cos \omega \tau d\tau \\ &= \frac{x_0^2}{2} \int_{-\infty}^{\infty} \cos(\omega - \omega_0)\tau d\tau + \frac{x_0^2}{2} \int_{-\infty}^{\infty} \cos(\omega + \omega_0)\tau d\tau \\ &= \frac{x_0^2}{2} (\delta(\omega - \omega_0) + \delta(\omega + \omega_0)) . \end{aligned}$$

Łatwa separacja częstotliwości:
 - dwa impulsy Diraca dla $-\omega_o$ i ω_o



Sygnały okresowe

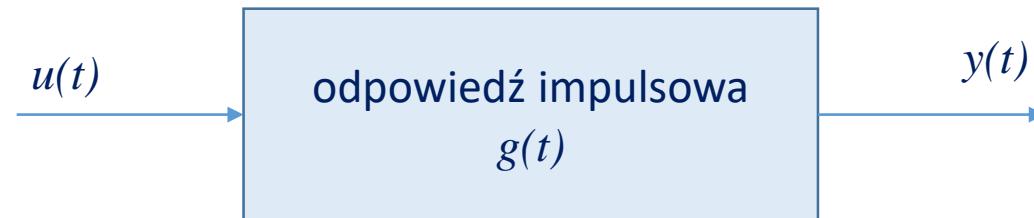
- Funkcja korelacji wzajemnej dla dwóch sygnałów:

- $x(t) = x_0 \sin(n\omega_0 t + \alpha_n)$ $n = 1, 2, 3, \dots$
- $y(t) = y_0 \sin(m\omega_0 t + \alpha_m)$ $m = 1, 2, 3, \dots$

$$R_{xy}(\tau) = \frac{x_0 y_0}{T_0} \int_0^{\frac{T}{2}} \sin(n\omega_0 t + \alpha_n) \sin(m\omega_0(t + \tau) + \alpha_m) dt = 0 \text{ if } n \neq m$$

- występują tylko wspólne częstotliwości

Procesy liniowe



- Korelacja wzajemna

$$R_{uy}(\tau) = E\{u(t)y(t + \tau)\}$$

$$\begin{aligned} R_{uy}(\tau) &= E\left\{u(t) \int_0^{\infty} g(t') u(t + \tau - t') dt'\right\} \\ &= \int_0^{\infty} g(t') E\{u(t)u(t + \tau - t')\} dt' \\ &= \int_0^{\infty} g(t') R_{uu}(\tau - t') dt' . \end{aligned}$$

Procesy liniowe

- Wzajemna gęstość spektralna (otrzymana przekształceniem Fouriera)

$$\begin{aligned}
 S_{uy}(i\omega) &= \int_{-\infty}^{\infty} R_{uy}(\tau) e^{-i\omega\tau} d\tau \\
 &= \int_{-\infty}^{\infty} \int_0^{\infty} g(t') R_{uu}(\tau - t') dt' e^{-i\omega\tau} d\tau \\
 &= \int_0^{\infty} g(t') dt' \int_{-\infty}^{\infty} R_{uu}(\tau - t') dt' e^{-i\omega\tau} d\tau \\
 &= \int_0^{\infty} g(t') e^{-i\omega t'} dt' S_{uu}(i\omega) .
 \end{aligned}$$

- Czyli otrzymujemy

$$S_{uy}(i\omega) = G(i\omega) S_{uu}(i\omega)$$

- I dalej:

$$S_{yy}(i\omega) = G(i\omega) S_{yu}(i\omega)$$

$$S_{yu}(i\omega) = S_{uy}(-i\omega)$$

$$S_{yy}(i\omega) = G(i\omega) G(-i\omega) S_{uu}(i\omega) = |G(i\omega)|^2 S_{uu}(i\omega)$$

zespolona transmitancja sprzężona

Układ liniowy

- Jeśli wejściem jest biały szum S_0 , to otrzymujemy na wyjściu szum kolorowy

$$S_{yy}(i\omega) = |G(i\omega)|^2 S_0$$

Procesy z czasem dyskretnym (stacjonarne)

- Średnia

$$\bar{x} = E\{x(k)\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N x(k)$$

- Kwadratowa średnia (wariancja)

$$\sigma_x^2 = E\{(x(k) - \bar{x})^2\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N (x(k) - \bar{x})^2$$

- Funkcja autokorelacji ACF

$$R_{xx}(\tau) = E\{x(k)x(k + \tau)\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N x(k)x(k + \tau)$$

- Funkcja korelacji wzajemnej CCF

$$R_{xy}(\tau) = E\{x(k)y(k + \tau)\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N x(k)y(k + \tau)$$

$$= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N x(k - \tau)y(k)$$

Procesy z czasem dyskretnym (stacjonarne)

- Funkcja autokowariancji

$$\begin{aligned} C_{xx}(\tau) &= \text{cov}(x, \tau) = E\{(x(k) - \bar{x})(x(k + \tau) - \bar{x})\} \\ &= E\{x(k)x(k + \tau)\} - \bar{x}^2 \end{aligned}$$

- Funkcja kowariancji wzajemnej

$$\begin{aligned} C_{xy}(\tau) &= \text{cov}(x, y, \tau) = E\{(x(k) - \bar{x})(y(k + \tau) - \bar{y})\} \\ &= E\{x(k)y(k + \tau)\} - \bar{x}\bar{y} \end{aligned}$$

Pamiętamy, iż kowariancja (korelacja) nic nie mówi o zależnościach przyczynowo-skutkowych

- Gęstość widmowa

$$S_{xx}^*(i\omega) = \mathfrak{F}\{R_{xx}(\tau)\} = \sum_{\tau=-\infty}^{\infty} R_{xx}(\tau) e^{-i\tau\omega T_0}$$

Procesy z czasem dyskretnym (stacjonarne)

- Szum biały

- Funkcja autokorelacji

$$R_{xx}(\tau) = \sigma_x^2 \delta(\tau)$$

- σ_x^2 - wariancja
 - z deltą Kroneckera

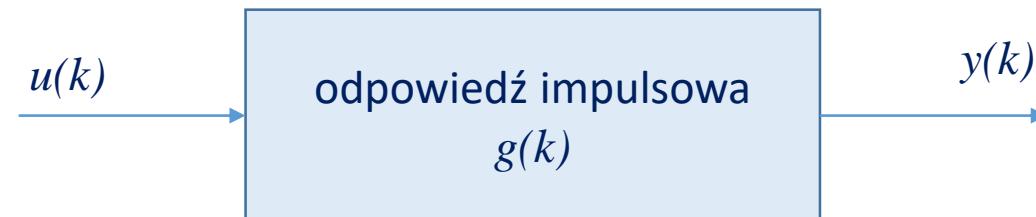
$$\delta(k) = \begin{cases} 1 & \text{for } k = 0 \\ 0 & \text{for } k \neq 0 \end{cases}$$

- Gęstość widmowa (stała na $0 \leq |\omega| \leq \pi/T_0$)

$$S_{xx}(z) = \sigma_x^2 \sum_{\tau=-\infty}^{\infty} \delta(\tau) z^{-\tau} = \sigma_x^2 = S_{xx0} = \text{const}$$

- Wariancja jest skończona, zatem w przeciwnieństwie do wersji ciągłej jest realizowalna

Proces liniowy z sygnałami dyskretnymi



- Funkcje korelacji wzajemnej i autokorelacji są powiązane analogicznie:

$$R_{uy}(\tau) = \sum_{k=0}^{\infty} g(k) R_{uu}(\tau - k)$$

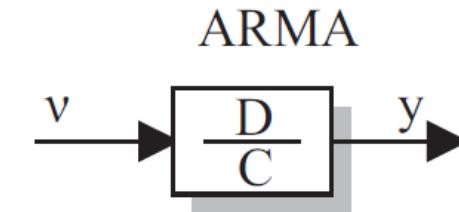
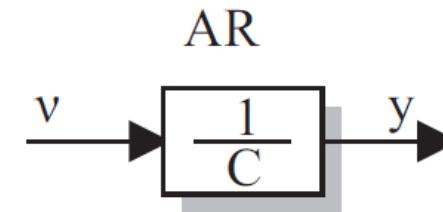
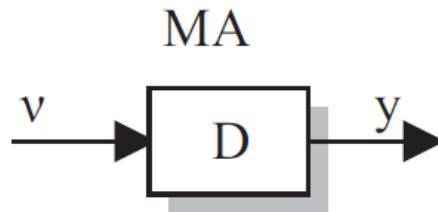
- A dla gęstości widmowej:

$$S_{uy}^*(i\omega) = G^*(i\omega) S_{uu}^*(i\omega) \text{ in the interval } |\omega| \leq \frac{\pi}{T_0}$$

$$S_{uy}(z) = G(z) S_{uu}(z)$$

Modele stochastyczne

- Modele procesów z sygnałami stochastycznymi



- Opisany stochastycznym równaniem różnicowym

$$y(k) + c_1 y(k-1) + \dots + c_n y(k-n) = d_0 v(k) + d_1 v(k-1) + \dots + d_m v(k-m)$$

- Z transmitancją filtra (modelu)

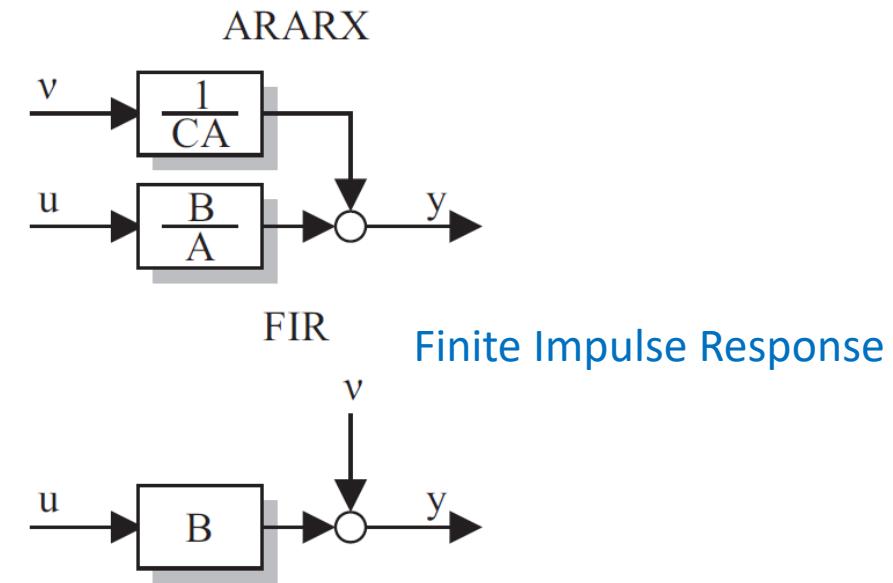
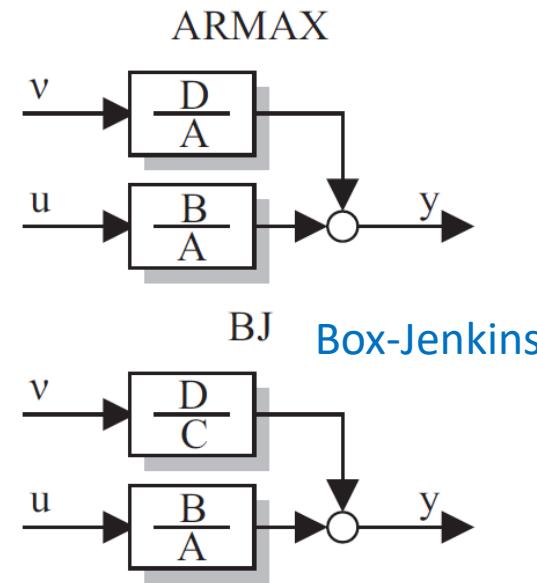
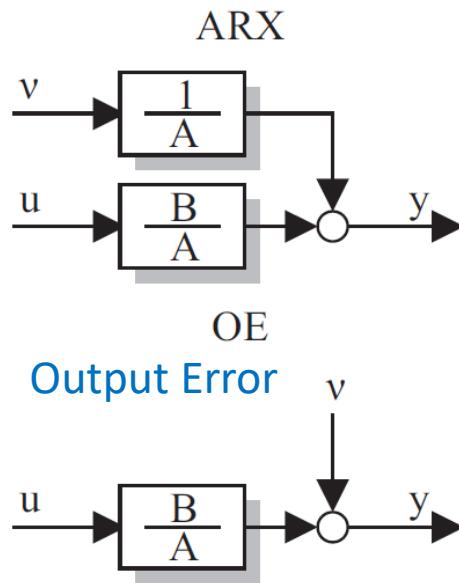
$$G_F(z^{-1}) = \frac{y(z)}{v(z)} = \frac{d_0 + d_1 z^{-1} + \dots + d_m z^{-m}}{1 + c_1 z^{-1} + \dots + c_n z^{-n}} = \frac{D(z^{-1})}{C(z^{-1})}$$

- AR: $y(k) + c_1 y(k-1) + \dots + c_n y(k-n) = d_0 v(k)$

- MA: $y(k) = d_0 v(k) + d_1 v(k-1) + \dots + d_m v(k-m)$

Modele stochastyczne

- Modele deterministyczne z zakłóceniami stochastycznymi (auxiliary, exogenous)



- Przykładowo ARX:

$$y(k) + c_1y(k-1) + \dots + c_ny(k-n) = d_0v(k) + b_1u(k-1) + \dots + b_nu(k-n)$$



MI

Metody Identyfikacji

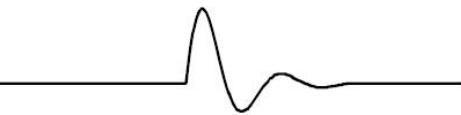
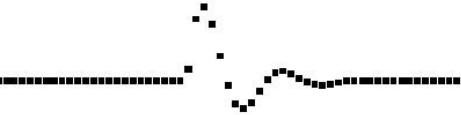
wykład #3

1. *Metody analizy spektralnej dla sygnałów okresowych i nieokresowych*

Analiza spektralna

- Wyznaczenie spektrum sygnałów jest często potrzebne
- Numeryczne sposoby wyliczenia transformaty Fouriera
- Tym samym potrzebne próbkowanie i metody cyfrowe
- Okienkowanie danych
- Efektywność obliczeniowa transformaty Fouriera dla dużych ilości danych

Transformaty

Type of Transform	Example Signal
Fourier Transform <i>signals that are continuous and aperiodic</i>	
Fourier Series <i>signals that are continuous and periodic</i>	
Discrete Time Fourier Transform <i>signals that are discrete and aperiodic</i>	
Discrete Fourier Transform <i>signals that are discrete and periodic</i>	

Numeryczne wyznaczanie transformaty Fouriera

- Cel: zawartość częstotliwościowa sygnału
- Wyznaczanie odpowiedzi częstotliwościowej modelu:
 - Konieczność wyznaczenia transformaty dla wejścia i wyjścia
- Wejście i wyjście w postaci dyskretyzowanego sygnału t_k dla $k = 0, 1, 2, \dots, N$
- Konieczność wypracowania **Dyskretnej Transformaty Fouriera**
- **Szybka Transformata Fouriera** znacznie bardziej efektywna obliczeniowo

Ciąg Fouriera dla sygnału okresowego

- Każdą funkcję okresową $x(t)$ z okresem T , tj. $x(t) = x(t+kT)$ można opisać jako

$$x(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(k\omega_0 t) + b_k \sin(k\omega_0 t) \text{ with } \omega_0 = \frac{2\pi}{T}$$

- Taki szereg jest nazywany **szeregiem Fouriera**.
- Uwzględniamy raczej skończoną liczbę elementów. Parametry opisane jako:

$$a_k = \frac{2}{T} \int_0^T x(t) \cos(k\omega_0 t) dt$$

$$b_k = \frac{2}{T} \int_0^T x(t) \sin(k\omega_0 t) dt$$

- W dziedzinie zespolonej otrzymujemy postać:

$$x(t) = \sum_{k=-\infty}^{\infty} c_k e^{ik\omega_0 t}$$

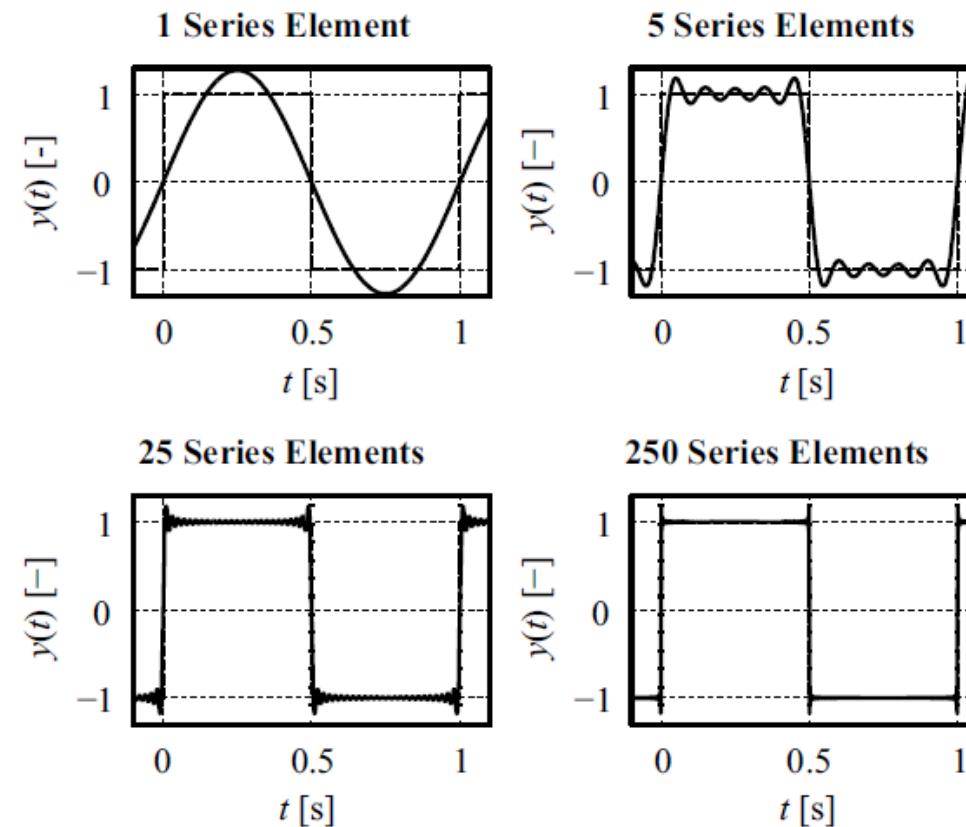
$$c_k = \frac{1}{T} \int_0^T x(t) e^{-ik\omega_0 t} dt$$

Ciąg Fouriera dla sygnału okresowego: efekt Gibbsa

- Efekt Gibbsa (1899) nazwany od imienia *Josiah Willarda Gibbsa* w matematyce jest to charakterystyczny sposób, w jaki zachowuje się aproksymacja funkcji $x(t)$ szeregiem Fouriera w punktach nieciągłości tej funkcji. Wykres nadmiernie oscyluje wokół tego punktu. Można przyjąć, że zjawisko to odzwierciedla trudność naśladowania nieciągłej funkcji przez skończone szeregi sinusów – nawet jeśli ilość elementów dąży do nieskończoności.
- Efekt Gibbsa wyjaśnia powstawanie zakłóceń różnego rodzaju sygnałów i znajduje zastosowanie w przetwarzaniu sygnałów (na przykład w cyfrowej obróbce obrazów). Między innymi wyjaśnia on przyczynę powstawania wysokoczęstotliwościowych oscylacji stanowiących zakłócenia sygnału przy zastosowaniu filtrów o prostokątnych oknach.

Ciąg Fouriera dla sygnału okresowego: efekt Gibbsa

- Zawsze pojawi się przeregulowanie, dla $N \rightarrow \infty$ otrzymamy około 18%



Ciąg Fouriera dla sygnału nieokresowego

- Rozszerzamy czasokres sygnału $T \rightarrow \infty$ i otrzymujemy transformatę Fouriera jako:

$$\mathcal{F}\{x(t)\} = x(i\omega) = \int_{-\infty}^{\infty} x(t)e^{-i\omega t} dt$$

- Jeśli nieokresowy sygnał ciągły jest próbkowany z czasem T_0 to sygnał ten możemy opisać jako ciąg delt Diraca o określonych wysokościach

$$x_\delta(k) = \sum_{k=-\infty}^{\infty} x(t) \delta(t - kT_0) = \sum_{k=-\infty}^{\infty} x(kT_0) \delta(t - kT_0)$$

- i otrzymujemy

$$x_\delta(i\omega) = \int_{-\infty}^{\infty} \sum_{k=-\infty}^{\infty} x(kT_0) \delta(t - kT_0) e^{-ik\omega T_0} dt = \sum_{k=-\infty}^{\infty} x(kT_0) e^{-ik\omega T_0}$$

- Ta transformata nazywana jest **Dyskretną Transformatą Fouriera** (DTFT – Discrete Time Fourier Transform)

Ciąg Fouriera dla sygnału nieokresowego

- Transformata odwrotna DTFT ma postać

$$x(k) = \frac{T_0}{2\pi} \int_{-\frac{\pi}{T_0}}^{\frac{\pi}{T_0}} x_\delta(i\omega) e^{ik\omega T_0} d\omega$$

- Jako że ciągły sygnał $x(t)$ jest przemnażany przez szereg delt Diraca, wynikowa transformata Fouriera $x_\delta(i\omega)$ jest okresowa w dziedzinie częstotliwości.
- Mnożenie w dziedzinie czasu jest splotem w dziedzinie częstotliwości.
- Ponieważ spektrum jest okresowe wystarczy je wyznaczyć w zakresie
 - $0 \leq \omega < 2\pi/T_0$
lub
 - $-\pi/T_0 \leq \omega < +\pi/T_0$

Ciąg Fouriera dla sygnału nieokresowego

- Okresowość uzasadnia twierdzenie Shannona-Kotielnikowa (1948), mówiące iż tylko częstotliwości do

$$f \leq \frac{1}{2T_0} = \frac{1}{2} f_s$$

- gdzie f_s jest częstotliwością próbkowania (częstotliwość Nyquista) mogą być próbowane poprawnie. Pozostałe częstotliwości zostaną odtworzone fałszywie – zjawisko maskowania (**aliasing**).

Ciąg Fouriera dla sygnału nieokresowego

- Ponieważ nie możemy sumować do nieskończoności $-\infty \leq k \leq +\infty$, a ilość punktów wynosi N ($0 \leq k \leq N-1$) otrzymujemy Dyskretną Transformatę Fouriera w postaci:

$$\begin{aligned}x(i\omega) &= \sum_{k=0}^{N-1} x(kT_0)e^{-ik\omega T_0} \\&= \sum_{k=0}^{N-1} x(kT_0) \cos(k\omega T_0) - i \sum_{k=0}^{N-1} x(kT_0) \sin(k\omega T_0) \\&= \operatorname{Re}\{x(i\omega)\} + \operatorname{Im}\{x(i\omega)\}.\end{aligned}$$

- Ograniczenie ilości próbek prowadzi do **okienkowania** (ang. windowing).

Ciąg Fouriera dla sygnału nieokresowego

- Pomimo że spektrum częstotliwościowe jest ciągłą i okresową funkcją niestety w obliczeniach nie możemy przechować wszystkich częstotliwości. Tym samym dyskretyzujemy również częstotliwość ω .
- Wystarczy to zrobić w zakresie $0 \leq \omega < 2\pi/T_0$
- Spektrum częstotliwości jest również przemnożone przez funkcję próbującą

$$\tilde{x}(iv\Delta\omega) = \sum_{v=0}^{M-1} x(i\omega)\delta(i\omega - iv\Delta\omega)$$

- gdzie:
 - $\tilde{x}(iv\Delta\omega)$ oznacza próbowaną transformatę,
 - $\Delta\omega$ – przyrost częstotliwości
 - M – ilość punktów ($M=2\pi/(T_0\Delta\omega)$)
- Prowadzi to do splotu w dziedzinie czasu i otrzymujemy okresowość w dziedzinie czasu poza próbowanym zakresem, tj.

$$x(kT_0) = x(kT_0 + \mu T_n), \quad \mu = 0, 1, 2, \dots, \text{ and } T_n = \frac{2\pi}{\Delta\omega}$$

Ciąg Fouriera dla sygnału nieokresowego

- Dobierając okres próbkowania częstotliwości jako $T_n = N$ to tym samym okresowość jest równoznaczna z długością pomiaru w dziedzinie czasu. Zatem $M = N$.
- Ostatecznie otrzymujemy parę transformat DFT jako:

$$x(in\Delta\omega) = \text{DFT}\{x(kT_0)\} = \sum_{k=0}^{N-1} x(kT_0)e^{-ikn\Delta\omega T_0}$$

$$x(kT_0) = \text{DFT}^{-1}\{x_s(in\Delta\omega)\} = \sum_{k=0}^{N-1} x(in\Delta\omega)e^{ikn\Delta\omega T_0}$$

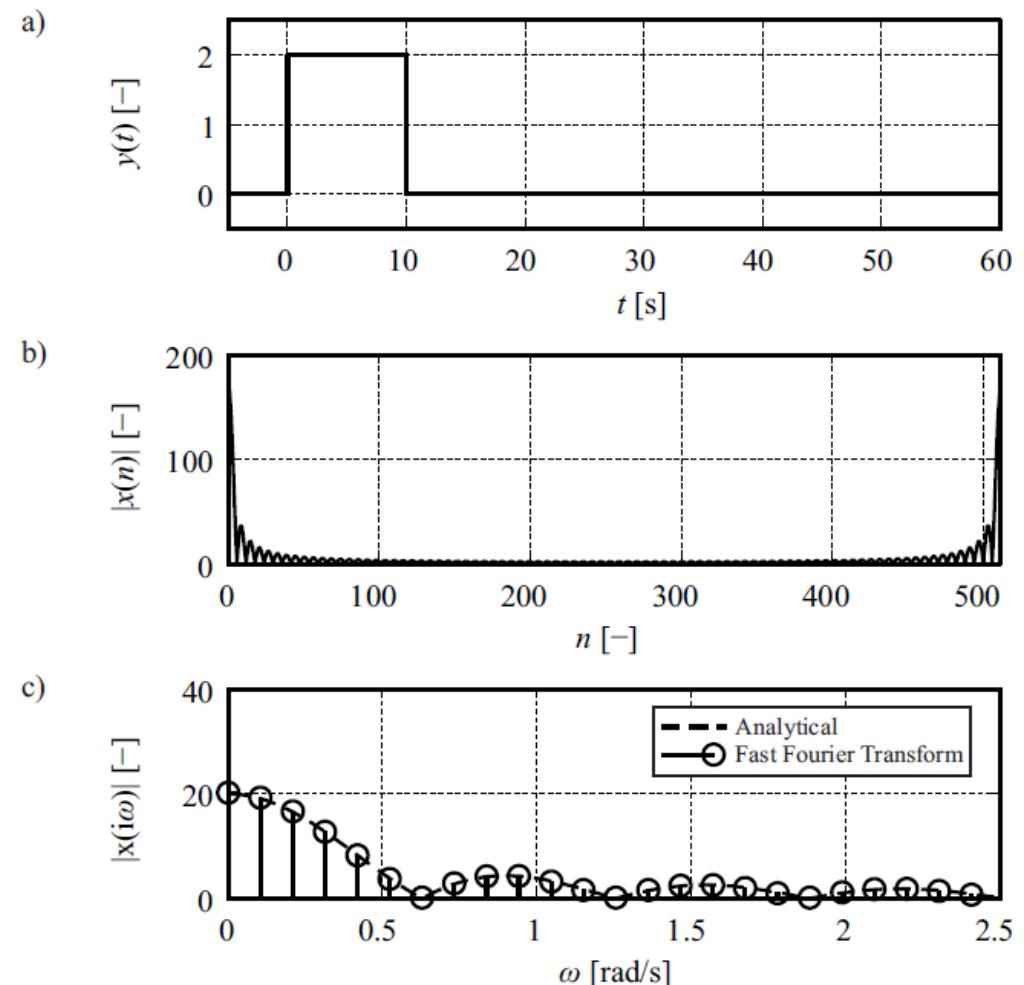
- Zastosowanie DFT powoduje, że zarówno sygnał jak i jego spektrum staje się okresowym.
- Dla każdej częstotliwości ω DFT potrzebuje N mnożeń i $(N-1)$ dodawań. Zatem dla całego spektrum potrzeba N^2 mnożeń i $N(N-1)$ dodawań. Wydatek obliczeniowy jest wysoki. Dlatego wprowadzimy **Szybką Transformację Fouriera (FFT – Fast Fourier Transform)**.

Ciąg Fouriera dla sygnału nieokresowego - Interpretacja DFT (FFT)

- FFT i analityczna postać transformaty Fouriera dla skoku prostokątnego:
 - Sygnal oryginalny
 - Wyjście FFT (nieskalowane)
 - Wyjście przeskalone zgodnie z:

$$\mathcal{F}\{x(kT_0)\} = T_0 x(i k \Delta \omega)$$

$$\omega = (0, \Delta\omega, 2\Delta\omega, \dots, (N-1)\Delta\omega) \text{ with } \Delta\omega = \frac{2\pi}{T_M}$$



Wprowadzenie teoretyczne

- Analiza Fouriera

Bardzo często wielkości mają charakter okresowy, tzn. taki, który powoduje powtarzanie się danej wielkości fizycznej z określonym okresem. Zazwyczaj taką funkcję okresową można przedstawić w postaci nieskończonego szeregu trygonometrycznego zwanego też szeregiem Fouriera.

$$f(x) = a_0 + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi x}{L} + b_n \sin \frac{n\pi x}{L} \right)$$

$$a_0 = \frac{1}{2L} \int_{-L}^{L} f(x) dx$$

$$a_n = \frac{1}{L} \int_{-L}^{L} f(x) \cos \frac{n\pi x}{L} dx$$

$$b_n = \frac{1}{L} \int_{-L}^{L} f(x) \sin \frac{n\pi x}{L} dx$$

Wprowadzenie teoretyczne

- Dyskretna transformata Fouriera
- Przypuśćmy, że mamy N kolejnych wartości zmierzonych w odstępach czasu Δ , tak że

$$h_k \equiv h(t_k), \quad t_k = k\Delta, \quad k = 0, 1, 2, \dots, N - 1$$

- Zamiast próbować znaleźć transformatę dla wszystkich wartości f oszacujmy ją jedynie w konkretnych punktach, danych przez:

$$f_n \equiv \frac{n}{N\Delta}, \quad n = -\frac{N}{2}, \dots, \frac{N}{2}$$

- Po przybliżeniu całki otrzymujemy $H_n \equiv \sum_{k=0}^{N-1} h_k e^{2\pi i kn/N}$
- Zastosowane powyżej przekształcenie nosi nazwę dyskretnej transformaty Fouriera

Algorytm FFT

- Uwagi ogólne
- Obliczanie transformaty bezpośrednio ze wzoru jest nieefektywne ze względu na zbyt dużą złożoność obliczeniową.
- Wzrost wydajności przy zastosowaniu FFT

	N	2	4	8	16	32	64	128
DFT	$2N^2$	8	32	128	512	2048	8192	32768
FFT	$\frac{1}{2} \log_2(N)$	2	8	24	64	160	384	896

Algorytm FFT

- Idea
- Sama idea algorytmu opiera się na tzw. lemacie Danielsona-Lanczosa. Odkryli oni, że pojedyncza DFT o długości N , jest równoważna sumie dwóch transformat o długości $N/2$, jedna z nich jest złożona z nieparzystych punktów spośród oryginalnych N , a druga z parzystych.

$$H_n = \sum_{j=0}^{N/2-1} e^{2\pi i n j / (N/2)} f_{2j} + W^n \sum_{j=0}^{N/2-1} e^{2\pi i n j / (N/2)} f_{2j+1} = \\ = H_n^e + W^n H_n^o$$

- H_n^e oznacza n -ty składnik transformaty o długości $N/2$, stworzony z parzystych (*even*) punktów, a H_n^o odpowiednio z nieparzystych (*odd*).

Algorytm FFT

- Algorytm Cooley'a-Tukey'a

Przykład wyznaczania transformaty dla $N = 8$ punktów											
		a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7		
		a_0	a_2	a_4	a_6			a_1	a_3	a_5	a_7
a_0	a_4			a_2	a_6	a_1	a_5			a_3	a_7
$000 = 0$	$100 = 4$	$010 = 2$	$110 = 6$	$001 = 1$	$101 = 5$	$011 = 3$	$111 = 7$				
a_0	a_4	a_2	a_6	a_1	a_5	a_3	a_7				
$000 = 0$	$001 = 1$	$010 = 2$	$011 = 3$	$100 = 4$	$101 = 5$	$110 = 6$	$111 = 7$				
$b_0 = a_0 + a_4$	$b_2 = a_2 + a_6$	$b_4 = a_1 + a_5$	$b_6 = a_3 + a_7$								
$b_1 = a_0 - a_4$	$b_3 = a_2 - a_6$	$b_5 = a_1 - a_5$	$b_7 = a_3 - a_7$								
$c_0 = b_0 + b_2$	$c_2 = b_0 - b_2$	$c_4 = b_4 + b_6$	$c_6 = b_4 - b_6$								
$c_1 = b_1 + \omega_4 b_3$	$c_3 = b_1 - \omega_4 b_3$	$c_5 = b_5 + \omega_4 b_7$	$c_7 = b_5 - \omega_4 b_7$								
		$d_0 = c_0 + c_4$	$d_4 = c_0 - c_4$								
		$d_1 = c_1 + \omega_8 c_5$	$d_5 = c_1 - \omega_8 c_5$								
		$d_2 = c_2 + \omega_8^2 c_6$	$d_6 = c_2 - \omega_8^2 c_6$								
		$d_3 = c_3 + \omega_8^3 c_7$	$d_7 = c_3 - \omega_8^3 c_7$								

$$W_N^{k-\frac{N}{2}} = e^{i2\pi(k-\frac{N}{2})/N} = e^{i2\pi k/N} e^{-i\pi} = -e^{i2\pi k/N} = -W_N^k$$

Zastosowanie analizy Fouriera

- Uwagi ogólne
- W ciągu ostatnich lat, wraz z rozwojem elektronicznej techniki obliczeniowej, nastąpił szybki rozwój teorii dotyczących analiz szeregów czasowych.
- Powstawały nowe metody numerycznego opracowania danych, które wcześniej nie mogły być zastosowane, ze względu na ogólną czasochłonność obliczeń.
- Metody te opracowywane były głównie dla potrzeb elektroniki gdzie, aby dostać np. dokładniejsze estymatory widm mocy lub lepszą filtrację, wydłużano szeregi czasowe.

Zastosowanie analizy Fouriera

- Analiza Fouriera w fizyce:

- Współczynniki Fouriera są interpretowane jako amplitudy odpowiednich składowych harmonicznych.
- Pierwsza składowa przekształcenia a_0 określa wartość stałą. Zależy ona od położenia sygnału względem osi poziomej. W praktyce jest najczęściej pomijana.
- Kwadraty współczynników z dokładnością do czynnika multiplikatywnego określają energię danej składowej harmonicznej.
- W ten sposób można mówić fizycznie o badaniu widma pewnej wielkości fizycznej tzn. rozkładzie energii w funkcji częstotliwości.

Zastosowanie analizy Fouriera

- Analiza Fouriera w elektronice:

- Widmo sygnału prostokątnego składa się z harmonicznych o częstotliwościach będących całkowitą nieparzystą wielokrotnością częstotliwości podstawowej o amplitudach malejących ze wzrostem częstotliwości harmonicznych.
- Im więcej składowych harmonicznych jest sumowanych tym lepsze jest przybliżenie przebiegu prostokątnego.
- W konkretnych zagadnieniach, kształt badanego sygnału jest na tyle skomplikowany, że trudno jest obliczyć go w sposób ścisły. Problemy z detekcją i szumami.
- Filtracja oraz pasmo przenoszenia sygnału.

Wyznaczanie numeryczne transformaty Fouriera

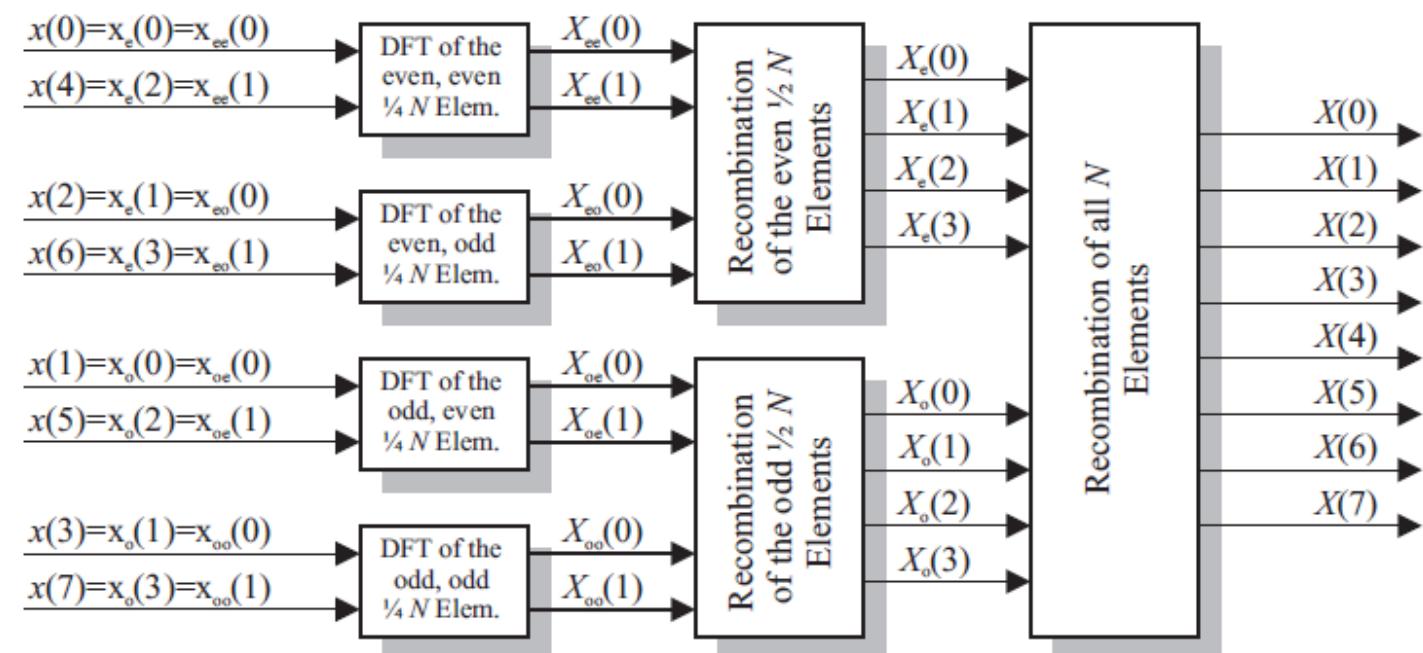
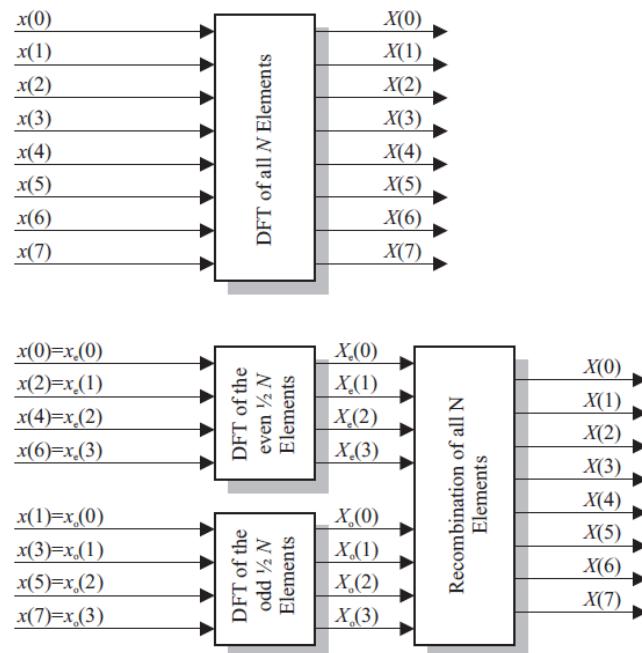
- Zakładając $e^{-ikn\Delta\omega T_0} = W_N^{nk}$ otrzymujemy $x(n) = \sum_{k=0}^{N-1} x(k)W_N^{nk}$
- Ze względu na właściwości operatora transformatę można podzielić na dwie sumy, dla parzystych i nieparzystych elementów:

$$\begin{aligned}
 x(n) &= \sum_{k=0}^{N-1} x(k)W_N^{nk} \\
 &= \sum_{k=0}^{\frac{N}{2}-1} x(2k)W_N^{2nk} + \sum_{k=0}^{\frac{N}{2}-1} x(2k+1)W_N^{n(2k+1)} \\
 &= \sum_{k=0}^{\frac{N}{2}-1} x(2k)W_N^{2nk} + W_N^n \sum_{k=0}^{\frac{N}{2}-1} x(2k+1)W_N^{2kn} \\
 &= x_e(n) + W_N^n x_o(n) ,
 \end{aligned}$$

DIT (radix-2 decimation-in-time) FFT
Algorytm Cooleya-Tukeya

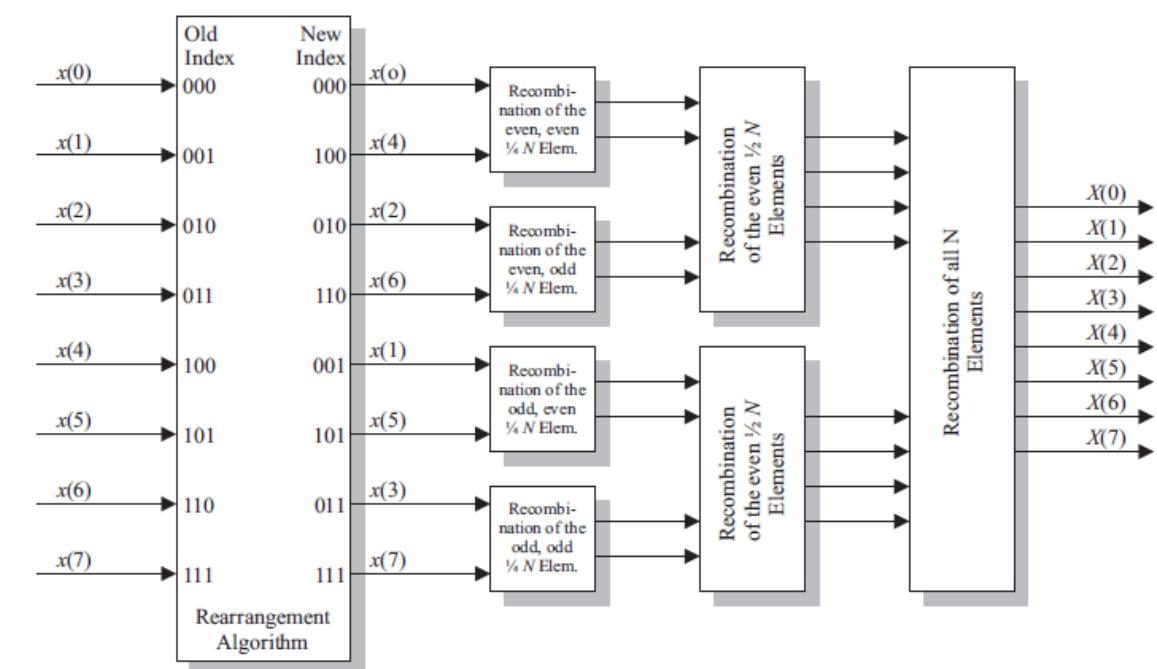
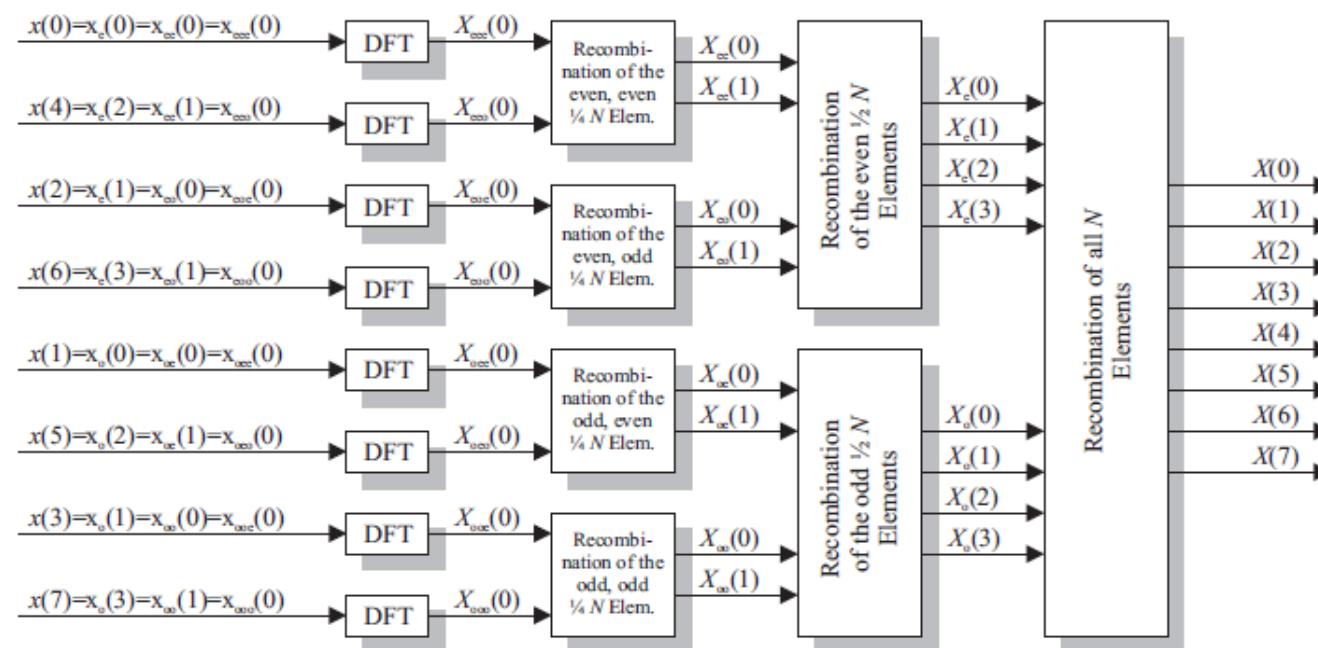
Wyznaczanie numeryczne transformaty Fouriera

- Obie sumy mają taki sam operator okresowy ($N/2$). Obie sumy są oznaczone „e” – parzysta i „o” – nieparzysta. Obie sumy można teraz podzielić wg qb na kolejne dwie dodając kolejny indeks, np. ”ee” parzysta-parzysta.
- Zasada dziel i rządź *divide et impera* dla $N = 8$ przedstawiona jest poniżej:



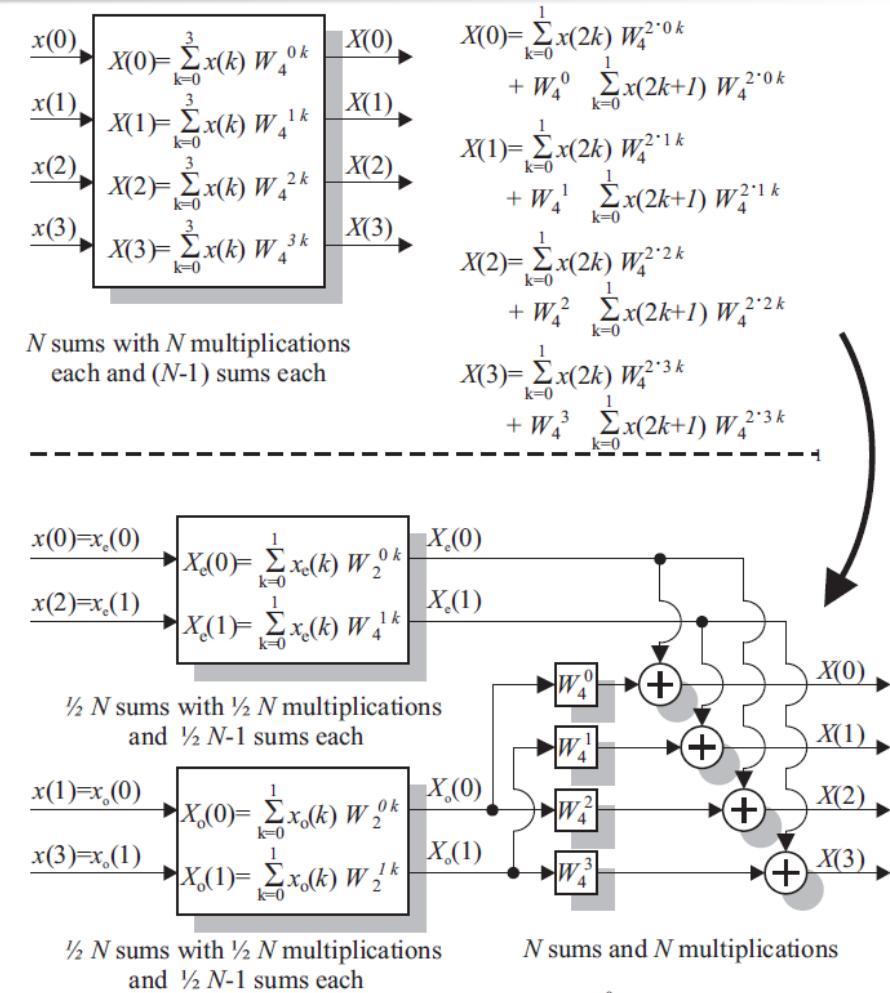
Wyznaczanie numeryczne transformaty Fouriera

- Zasada *divide et impera* dla $N = 8$ (ciąg dalszy) – rekombinacja:



Wyznaczanie numeryczne transformaty Fouriera

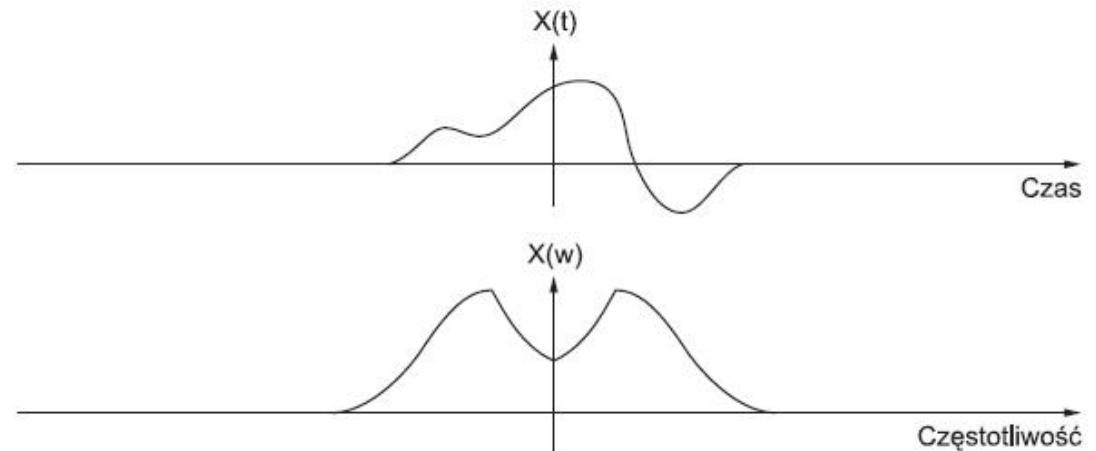
- Przed pierwszą rekombinacją należy przestawić próbki w odwrotnym porządku binarnym, np. element $n = 100_2$, stanie się elementem $n = 001_2$
- DFT potrzebuje N^2 zespolonych mnożeń i $N(N-1)$ zespolonych dodawań, podczas gdy $O(N)$ operacji może być oszczędzonych jako trywialne (mnożenie przez 1)
- Powyższy FFT ma złożoność $N \log_2 N$ zespolonych mnożeń i $N \log_2 N$ zespolonych dodawań, przy czym część z nich jest trywialna



$$\begin{aligned} X(0) &= X_e(0) + W_4^0 X_o(0) \\ X(1) &= X_e(1) + W_4^1 X_o(1) \\ X(2) &= X_e(2) + W_4^2 X_o(2) = X_e(0) + W_4^2 X_o(0) \\ X(3) &= X_e(3) + W_4^3 X_o(3) = X_e(1) + W_4^3 X_o(1) \end{aligned}$$

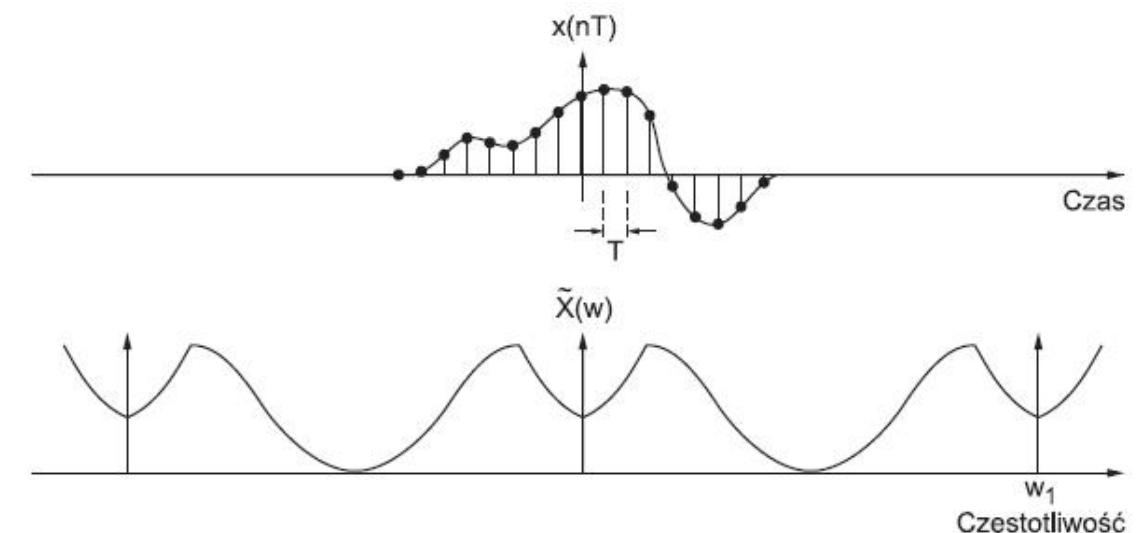
FFT – uwagi praktyczne

- Po pierwsze w praktyce ma się do czynienia z sygnałami o skończonym czasie trwania i z widmami o skończonej szerokości pasma.
- Po drugie, próbki sygnału muszą mieć określoną długość, a widmo określone częstotliwości składowe.
- Przykładowy ciągły w czasie nieokresowy $x(t)$ jest pokazany obok. Jego fourierowska transformata $X(\omega)$ jest ciągła i nieokresowa w domenie częstotliwości. Sygnał próbkujesz się w odstępach T , czyniąc go nieciągły w czasie.
- Pamiętamy, że mnożenie w domenie czasu odpowiada splotowi w domenie częstotliwości i na odwrót.



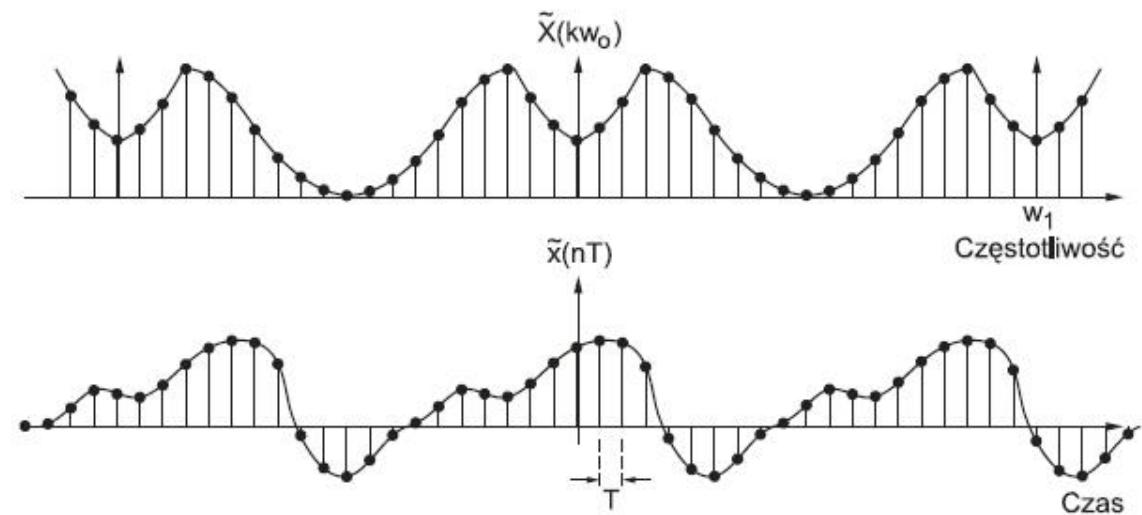
FFT – uwagi praktyczne

- Ponieważ próbkowanie jest mnożeniem $x(t)$ i sygnału próbkującego $s(t)$ w domenie czasu, odpowiada ono splotowi widma sygnału z widmem $s(t)$, będącego szeregiem harmonicznych f_s , pokazanym obok.
- Jeśli pamięta się proces i widmo modulacji amplitudy to można rozpoznać, że widmo jest widmem modulacji każdej harmonicznej f_s przez $s(t)$. Splot jest właśnie tym procesem.



FFT – uwagi praktyczne

- Następnym krokiem jest przemiana tego widma w nieokresowe przez próbkowanie z częstotliwością ω_0 Hz. Zabieg ten w domenie częstotliwości wpływa na sygnał w domenie czasu czyniąc go okresowym.
- Jest to wynik splotu sygnału z ciągiem impulsów próbujących o okresie $1/f_0$.
- Powstał w ten sposób okresowy w czasie sygnał i jego okresowa transformata fourierowska. Rozciągają się one do nieskończoności, ale są okresowe.

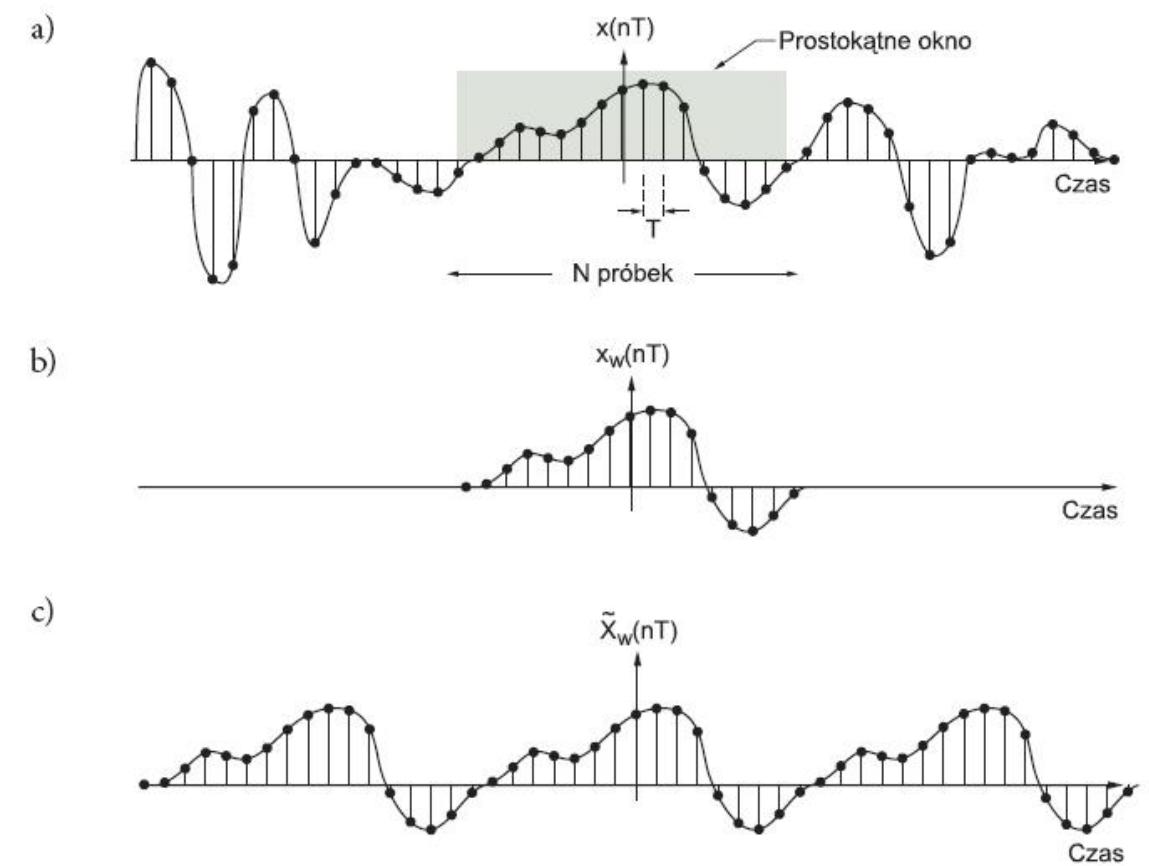


FFT – okienkowanie

- Trzeba więc wziąć jeden okres w każdej z domen mnożąc go przez funkcję prostokątną. Zabieg ten nazywa się okienkowaniem.
- Okienkowanie w jednej domenie wpływa oczywiście na sygnał w drugiej domenie.
- Jeden okres otrzymanego widma nieokresowego przyjmuje się więc za nieokresową fourierowską transformatę jednego okresu nieciągłego sygnału okresowego.

FFT – okienkowanie

- W praktyce do analizy sygnałów
 - najpierw trzeba przez próbkowanie otrzymać sygnał nieciągły, jak pokazano na (a),
 - wydzielić segment N próbek przez okienkowanie, jak pokazano na (b).
 - Następnie przyjmuje się, że segment jest powtarzalny, czyli segment zostaje powielony w obie strony, jak pokazuje (c).
- Dla tego sygnału oblicza się współczynniki jednego okresu nieciągłej transformaty Fouriera
- Następnie przenosi się na następny segment ciągu próbek i powtarza się obliczenia.
- Przedstawienie poprzez próbki transformat Fouriera jest w gruncie rzeczy przedstawieniem poprzez ciąg okresowy ciągu o ograniczonej długości, którego jeden okres jest skończoną sekwencją, która ma zostać przedstawiona.

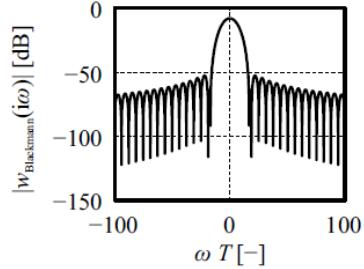
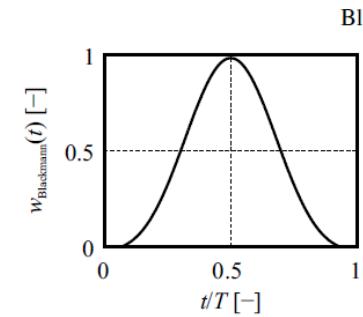
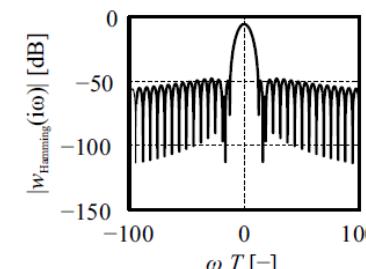
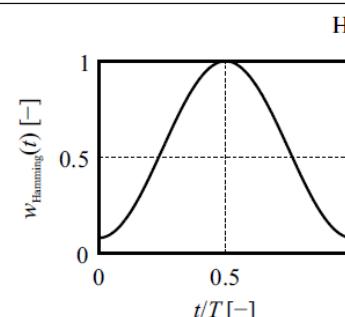
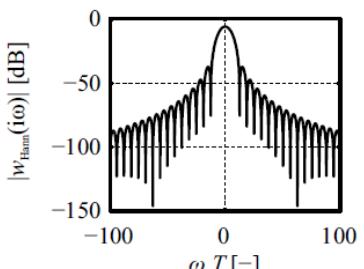
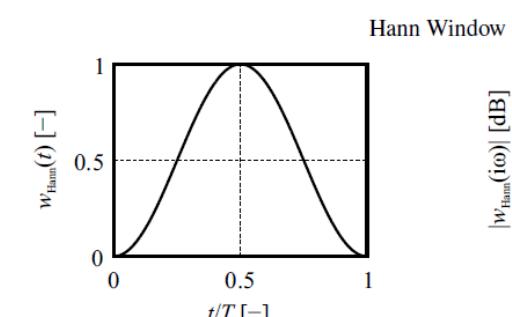
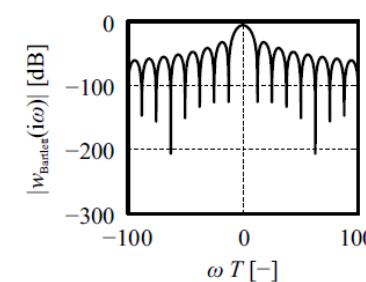
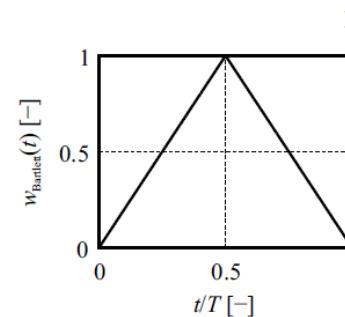


FFT – okienkowanie

- Różne rodzaje okien:

- Prostokątne (żadne) – największa ilość wycieku okna
- Hanninga – użyteczne do analizy przebiegów nieustalonych dłuższych niż czas trwania okna
- Hamminga - podobne do Hanninga
- Kaiser-Bessela – do wykrywania dwóch sygnałów o zbliżonej częstotliwości ale różnych amplitudach
- Trójkątne
- Flat Top – najwyższa dokładność amplitudowa kosztem selektywności częstotliwości – mierzenie amplitud pojedynczych składowych
- Wykładnicze – użyteczne przy mierzeniu sygnałów nieustalonych, tj. krótkotrwałych
- Bartlett
- Hann
- Blackmann

FFT – okienkowanie przykład



FFT – okienkowanie

Typ sygnału	Okno
Przebiegi nieustalone o czasie trwania krótszym od okna	Prostokątne
Przebiegi nieustalone o czasie trwania dłuższym od okna	Wykładnicze, Hanning
Aplikacje ogólnego zastosowania	Hanning
Śledzenie rozkazów	Prostokątne
Pomiary odpowiedzi częstotliwościowych	Hanning
Separacja tonów o zbliżonej częstotliwości ale różnych amplitudach	Kaiser-Bessel
Separacja tonów o różnej częstotliwości ale zbliżonych amplitudach	Prostokątne
Dokładny pomiar danej amplitudy	Flat Top

Periodogram

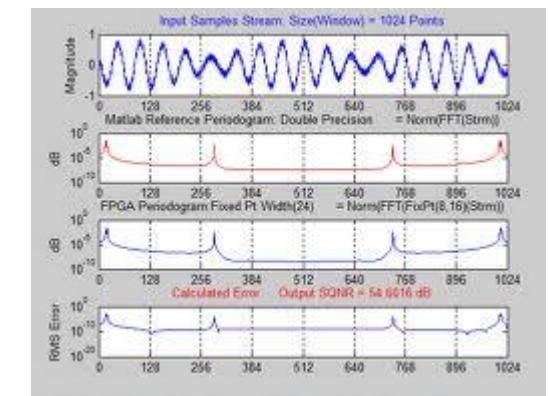
- Periodogram - rodzaj dyskretnej transformaty Fouriera. Pojęcia prawdopodobnie po raz pierwszy użył Arthur Schuster (1898), opierając się na pracy *Power Spectral Density estimation* (ang. "Estymacja widmowej gęstości mocy") Fernanda Schlindweina.

$$\hat{S}_{xx}(i\omega) = \frac{1}{N} |x(i\omega)|^2 = \frac{1}{N} x(i\omega)x^*(i\omega) = \frac{1}{N} \sum_{v=0}^{N-1} \sum_{\mu=0}^{N-1} x(v)x(\mu)e^{-i\omega(v+\mu)T_0}$$

- Narzędzie do wyznaczania spektrum sygnału

$$E\{\hat{S}_{xx}(i\omega)\} = \sum_{v=-(N-1)}^{N-1} w_{\text{Bartlett}}(v) R_{xx}(v) e^{-i\omega v T_0}$$

- Wartość oczekiwana jest dana jako rzeczywista gęstość widmowa spleciona z transformatą widmową okna Barletta.
- Jest tylko asymptotycznie nieobciążony a wariancja nie zbiera do zera jak $N \rightarrow \infty$.





MI

Metody Identyfikacji

wykład #4

1. *Wyznaczanie odpowiedzi częstotliwościowych dla sygnałów nieokresowych*

Periodogram

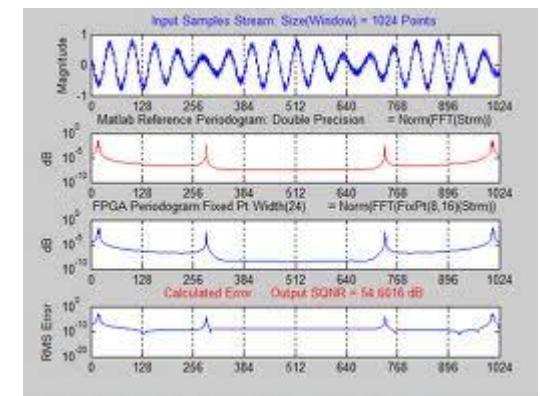
- Periodogram - rodzaj dyskretnej transformaty Fouriera. Pojęcia prawdopodobnie po raz pierwszy użył Arthur Schuster (1898), opierając się na pracy *Power Spectral Density estimation* (ang. "Estymacja widmowej gęstości mocy") Fernanda Schlindweina.

$$\hat{S}_{xx}(i\omega) = \frac{1}{N} |x(i\omega)|^2 = \frac{1}{N} x(i\omega)x^*(i\omega) = \frac{1}{N} \sum_{v=0}^{N-1} \sum_{\mu=0}^{N-1} x(v)x(\mu)e^{-i\omega(v+\mu)T_0}$$

- Narzędzie do wyznaczania spektrum sygnału

$$E\{\hat{S}_{xx}(i\omega)\} = \sum_{v=-(N-1)}^{N-1} w_{\text{Bartlett}}(v) R_{xx}(v) e^{-i\omega v T_0}$$

- Wartość oczekiwana jest dana jako rzeczywista gęstość widmowa spleciona z transformatą widmową okna Bartletta.
- Jest tylko asymptotycznie nieobciążony a wariancja nie zbiera do zera jak $N \rightarrow \infty$.



Analiza spektralna

- Nieparametryczna odpowiedź częstotliwościowa może być wyznaczona dla sygnałów nieokresowych z:

$$G(i\omega) = \frac{y(i\omega)}{u(i\omega)} = \frac{\mathfrak{F}\{y(t)\}}{\mathfrak{F}\{u(t)\}} = \frac{\int_0^\infty y(t)e^{-i\omega t} dt}{\int_0^\infty u(t)e^{-i\omega t} dt}$$

- gdzie całkę można rozbić na część rzeczywistą i urojoną:

$$y(i\omega) = \lim_{T \rightarrow \infty} \left(\int_0^T y(t) \cos \omega t dt - i \int_0^T y(t) \sin \omega t dt \right)$$

- tym samym musimy wyznaczyć transformatę Fouriera dla wejścia i wyjścia

Analiza spektralna

- Jeśli mamy do czynienia z nieokresowymi sygnałami to wtedy używamy transformaty Fouriera. I wtedy możemy użyć przekształcenia Laplace'a $s \rightarrow i\omega$:

$$G(i\omega) = \lim_{s \rightarrow i\omega} \frac{y(s)}{u(s)} = \lim_{s \rightarrow i\omega} \frac{\int_0^\infty y(t)e^{-st} dt}{\int_0^\infty u(t)e^{-st} dt} = \frac{y(i\omega)}{u(i\omega)}$$

- Analogicznie otrzymujemy część rzeczywistą i urojoną – i tak samo dla $u(i\omega)$:

$$y(i\omega) = \lim_{\substack{\delta \rightarrow 0 \\ T \rightarrow \infty}} \left(\int_0^T y(t)e^{-\delta t} \cos \omega t dt - i \int_0^T y(t)e^{-\delta t} \sin \omega t dt \right)$$

- W praktyce najczęściej operujemy odchyleniami od stanu ustalonego:

$$y(t) = Y(t) - Y_{00}$$

$$u(t) = U(t) - U_{00}$$

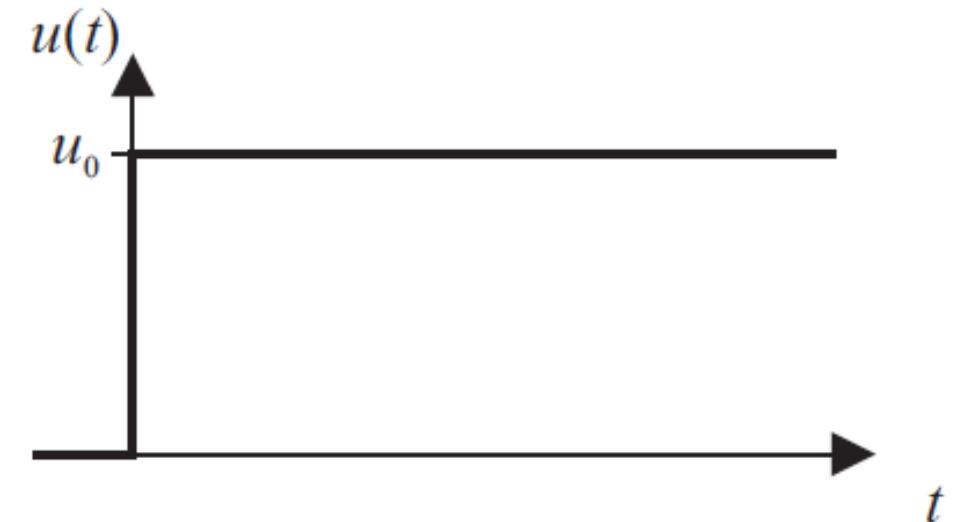
Analiza spektralna

- W praktyce stosujemy predefiniowane sygnały wejściowe, dla których znamy transformaty Fouriera
- Sygnał ten dobieramy w zależności od możliwości obiektu
- Wyznaczamy wtedy transformatę Fouriera jedynie dla wyjścia $y(t)$
- Najprostsze podejście – otrzymujemy pierwszy wgląd w model

Transformaty Fouriera przykładowych sygnałów

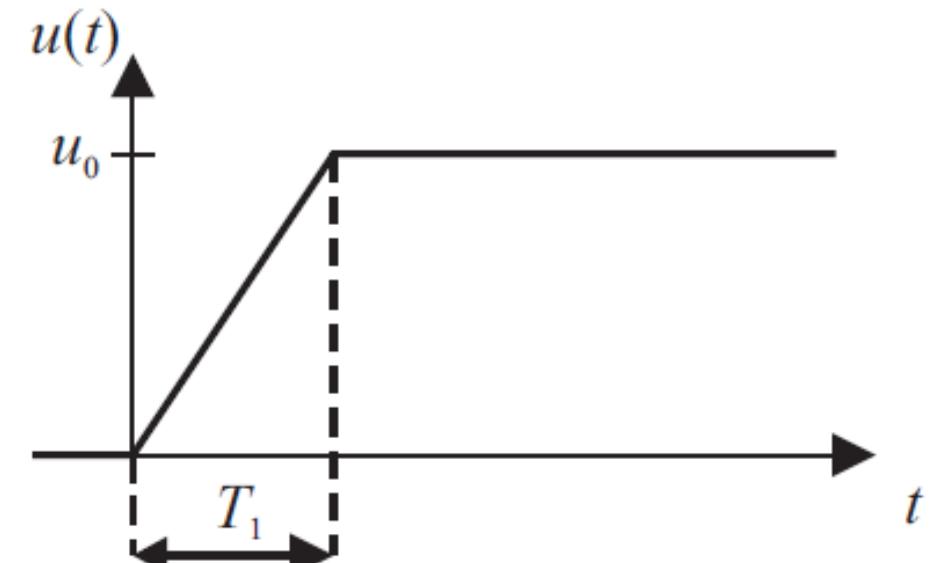
- Skok jednostkowy

- Zawsze dodatni
- Niedostosowany do obiektów z całkowaniem
- Nie można bezpośrednio wyznaczyć transformaty Fouriera



- „Rampa”

- Zawsze dodatni
- Niedostosowany do obiektów z całkowaniem
- Nie można bezpośrednio wyznaczyć transformaty Fouriera



Transformaty Fouriera przykładowych sygnałów

- Impuls trapezowy

- Można wyznaczyć transformatę Fouriera, gdy $T_1 = T_2$

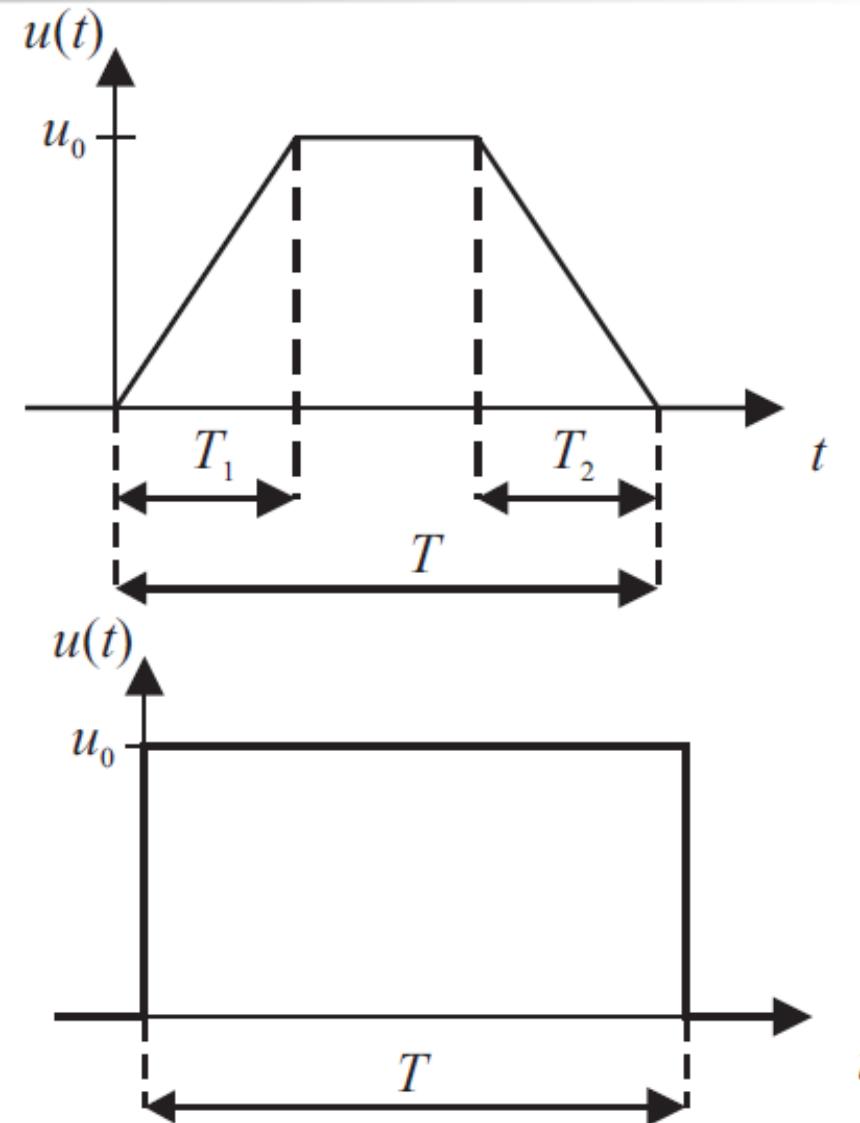
$$u_{\text{tr}}(i\omega) = u_0(T - T_1) \left(\frac{\sin \frac{\omega T_1}{2}}{\frac{\omega T_1}{2}} \right) \left(\frac{\sin \frac{\omega(T-T_1)}{2}}{\frac{\omega(T-T_1)}{2}} \right) e^{-i\frac{\omega T}{2}}$$

- Podwójny impuls prostokątny (otrzymujemy przyjmując, że $T_1 \rightarrow 0$)

$$u_{\text{sq}}(i\omega) = u_0 T \left(\frac{\sin \frac{\omega T}{2}}{\frac{\omega T}{2}} \right) e^{-i\frac{\omega T}{2}}$$

- Zalecenie (Pintelon i Schoukens) dla T : okresu T

$$T = \frac{1}{2.5f_{\max}}$$



Transformaty Fouriera przykładowych sygnałów

- Impuls trójkątny

- $T_1 = T_2 = T/2$

$$u_{\text{tri}}(i\omega) = u_0 \frac{T}{2} \left(\frac{\sin \frac{\omega T}{4}}{\frac{\omega T}{4}} \right)^2 e^{-i\frac{\omega T}{2}}$$

Transformaty Fouriera przykładowych sygnałów

- Reprezentacja bezwymiarowa:
 - $u^*(t) = u(t) / u_o$
 - $t^* = t / T$
 - $\omega^* = (\omega T) / 2\pi$
- Transformata Fouriera jest znormalizowana względem największej wartości osiąganej przez impuls prostokątny

$$u_{\text{sq}}(i\omega)|_{\omega=0} = \int_0^T u_0 dt = u_0 T$$

- Przy takim wyborze sygnały testowe o zbliżonym kształcie, ale różnych amplitudach o długościach impulsu mają identyczny moduł $|u^*(i\omega^*)|$ i fazę $< u^*(i\omega^*)$

Transformaty Fouriera przykładowych sygnałów

- Tym samym tylko kształt impulsu określa transformatę Fouriera

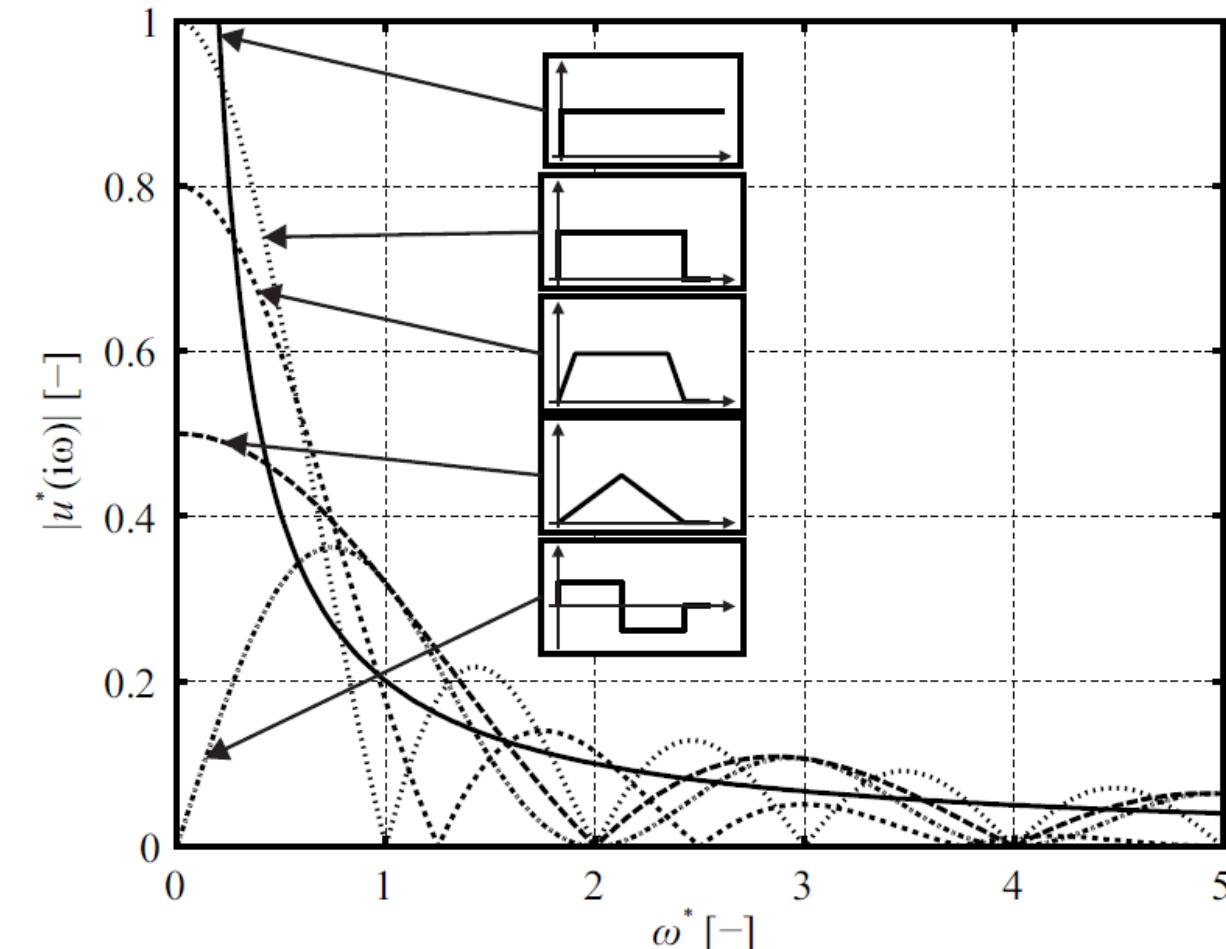
$$u_{\text{tr}}^*(i\omega^*) = (T^* - T_1^*) \left(\frac{\sin \pi \omega^* T_1^*}{\pi \omega^* T_1^*} \right) \left(\frac{\sin \pi \omega^* (T^* - T_1^*)}{\pi \omega^* (T^* - T_1^*)} \right) e^{-i\pi \omega^*}$$

$$u_{\text{sq}}^*(i\omega^*) = \left(\frac{\sin \pi \omega^*}{\pi \omega^*} \right) e^{-i\pi \omega^*}$$

$$u_{\text{tri}}^*(i\omega^*) = \frac{1}{2} \left(\frac{\sin \frac{\pi \omega^*}{2}}{\frac{\pi \omega^*}{2}} \right)^2 e^{-i\pi \omega^*}.$$

Transformaty Fouriera przykładowych sygnałów

- Amplitudy maleją aż do pierwszego zera wraz ze wzrostem częstotliwości
- Potem następują kolejne zera i pośrednie lokalne maksima
 - $\omega^* = n/T_1^*$ – pierwszy rząd zer trapezowy
 - $\omega^* = n/(T^* - T_1^*)$ – drugi rząd zer trapezowy
 - $\omega^* = n$ – zera prostokątny
 - $\omega^* = 2n$ – zera trójkątny
- Trapezowy i prostokątny mają pojedyncze zera i jest prostopadły do osi
- Trójkątny ma podwójne zera i jest styczny do osi



Transformaty Fouriera przykładowych sygnałów

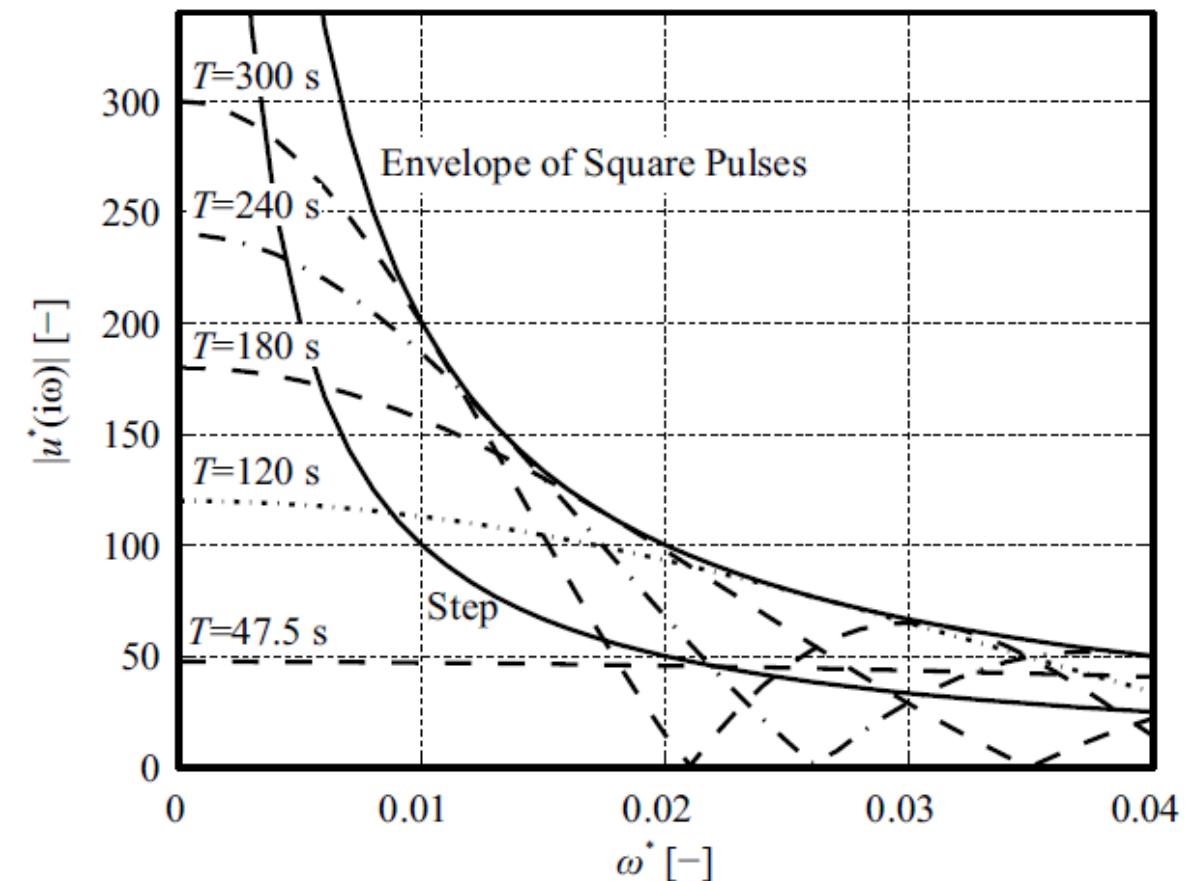
- $T \nearrow$, rośnie moduł dla niższych częstotliwości, rośnie pole pod impulsem
- Również zanikanie staje się szybsze dla wyższych częstotliwości
- Obwiednia (ang. envelope) oznacza najwyższą osiąganą amplitudę dla danej częstotliwości ω .

- impuls prostokątny

$$|u_{\text{sq}}^*(i\omega^*)|_{\max} = \frac{1}{\pi\omega^*} = \frac{0.3183}{\omega^*}$$

- impuls trójkątny

$$|u_{\text{tri}}^*(i\omega^*)|_{\max} = \frac{0.2302}{\omega^*}$$



Transformaty Fouriera przykładowych sygnałów

- Dla niskich częstotliwości pole pod impulsem determinuje amplitudę – impuls prostokątny ma największe pole dla danego okresu T.
- Dla średnich częstotliwości obwiednia wyznacza moduł. Impuls prostokątny ma najwyższą obwiednię w okolicy $\omega^* = 1/2$ a tym samym najwyższy moduł.
- Impuls prostokątny ma najwyższą obwiednię w całym zakresie częstotliwości $0 \leq \omega^* \leq 1/2$
- Dla wyższych jest to tylko zjawisko miejscowe

Transformaty Fouriera przykładowych sygnałów

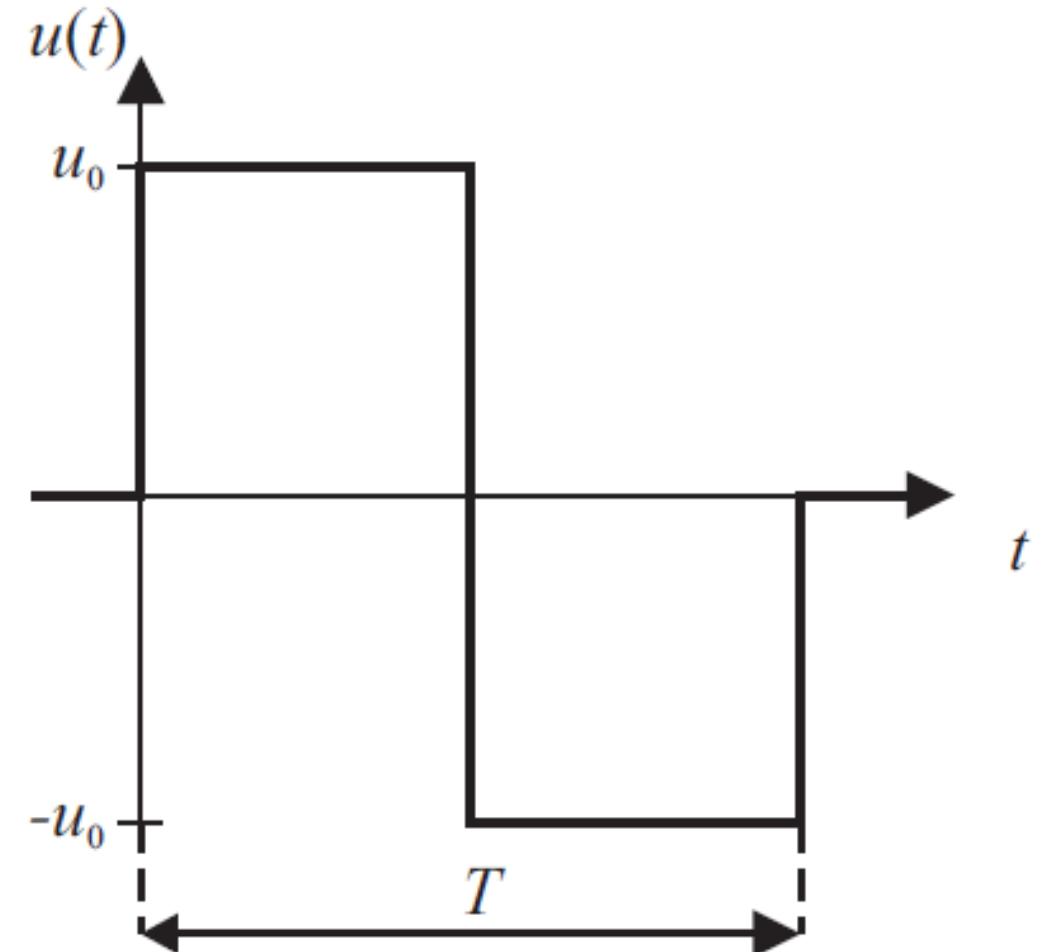
- Podwójny impuls prostokątny

$$u(i\omega) = u_0 T \left(\frac{\sin^2 \frac{\omega T}{4}}{\frac{\omega T}{4}} \right) e^{-i\frac{\omega T - \pi}{2}}$$

$$u^*(i\omega^*) = \left(\frac{\sin^2 \frac{\pi\omega^*}{2}}{\frac{\pi\omega^*}{2}} \right) e^{-i\pi \frac{2\omega^* - 1}{2}}$$

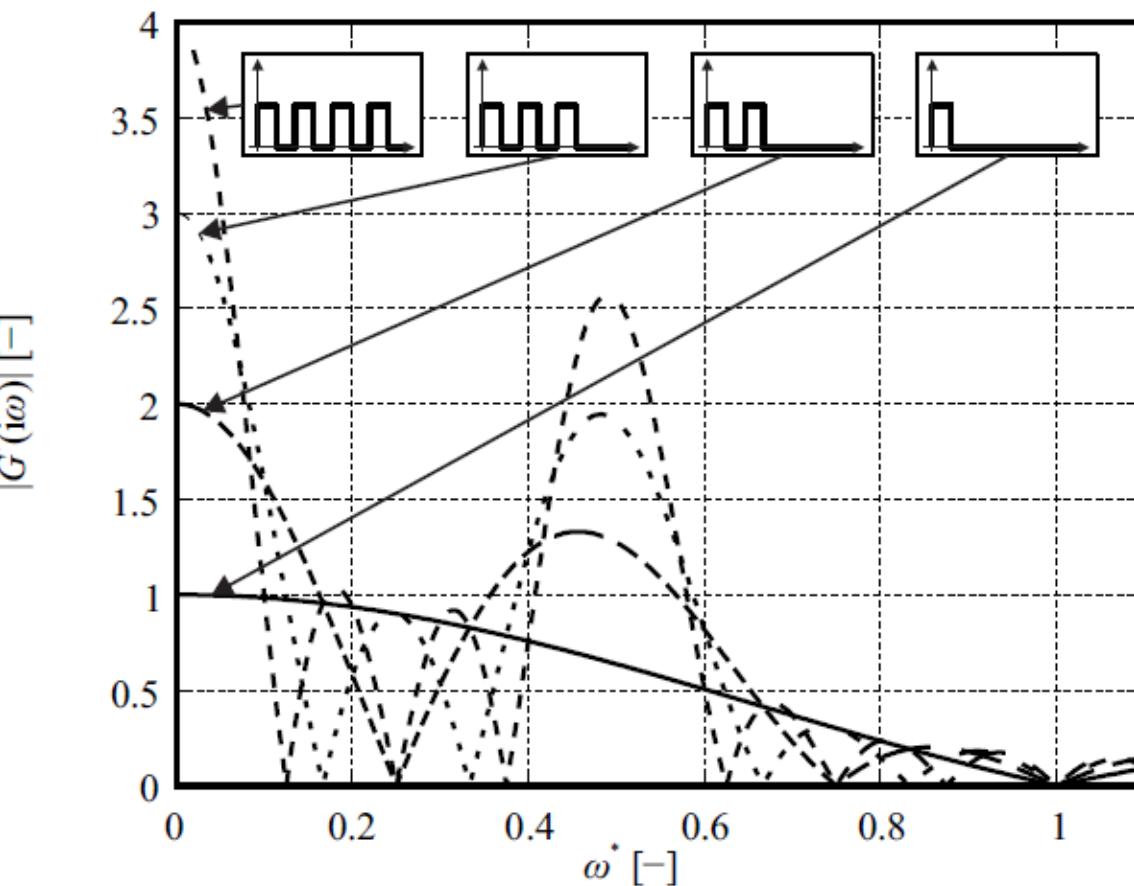
- Zera są dla $\omega^* = 2n$
- Zera są podwójne z wyjątkiem $n = 0$.
- Moduł wynosi zero dla $\omega^* = 0$
- Maksimum wynosi:

$$|u^*(i\omega^*)|_{\max} = 0.362 \text{ at } \omega^* = 0.762$$



Transformaty Fouriera przykładowych sygnałów

- Zmiany przebiegu transformaty Fouriera dla wzrostu ilości impulsów
 - Gdy otrzymamy ich nieskończoną liczbę będziemy mieli impulsy Diraca dla częstotliwości $0, 0.5, 1.5, 2.5, \dots$



Transformaty Fouriera przykładowych sygnałów

- Podsumowanie

- Im stromsze brzegi impulsu, tym większe wzbudzenie dla wyższych częstotliwości
- Pytanie czy można przyjąć skok zamiast „rampy”?
 - Przy akceptacji błędu $\leq 1\%$ otrzymamy okres wzrostu rampy $T_{1,max} \leq 1.1 / \omega_{max}$
 - Przy akceptacji błędu $\leq 5\%$ otrzymamy okres wzrostu rampy $T_{1,max} \leq 0.5 / \omega_{max}$
- Dla danej amplitudy impulsu u_o , najwyższa gęstość amplitudowa jest osiągana dla:
 - Skok jednostkowy dla niskich częstotliwości
 - Impulsy prostokątne dla średnich i wyższych częstotliwości
- Tym samym te sygnały zapewniają najmniejszy błąd identyfikacji charakterystyk częstotliwościowych dla zaszumionych wyjść
- Sygnały nieokresowe w przeciwnieństwie do okresowych pobudzają jednocześnie wszystkie częstotliwości za wyjątkiem zer dla impulsów oraz rampy

Wyznaczenie odpowiedzi częstotliwościowej

- Odpowiedź ma postać:

$$\hat{G}(i\omega) = \frac{y(i\omega)}{u(i\omega)} = \frac{\mathcal{F}\{y(t)\}}{\mathcal{F}\{u(t)\}}$$

- Zakładamy szum na wyjściu:

$$y(t) = y_u(t) + n(t)$$

- Po podstawieniu otrzymujemy:

$$\hat{G}(i\omega) = \frac{1}{u(i\omega)} \lim_{s \rightarrow i\omega} \left(\int_0^\infty y_u(t) e^{-st} dt + \int_0^\infty n(t) e^{-st} dt \right)$$

$$\hat{G}(i\omega) = G_0(i\omega) + \Delta G_n(i\omega)$$

- Zatem mamy złożenie odpowiedzi obiektu oraz charakterystyki szumu

$$\Delta G_n(i\omega) = \lim_{s \rightarrow i\omega} \frac{n(s)}{u(s)} = \frac{n(i\omega)}{u(i\omega)}$$

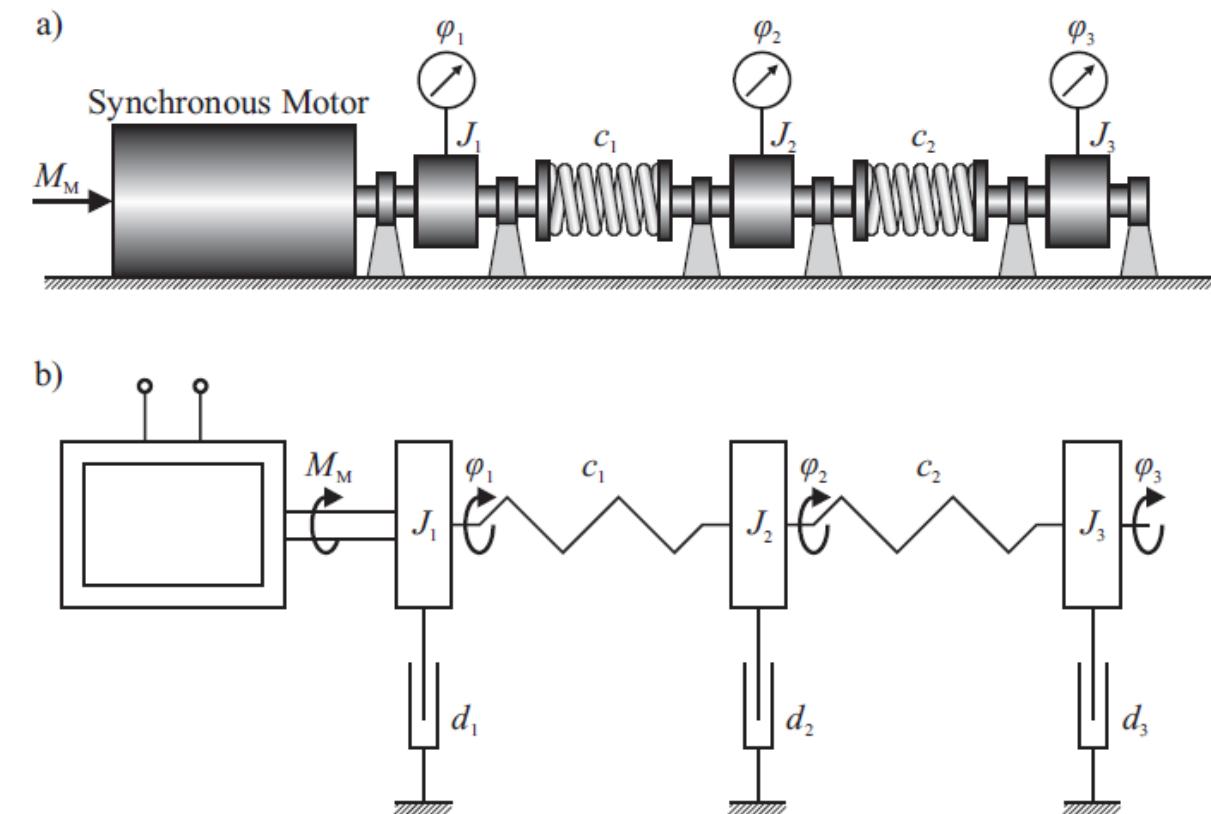
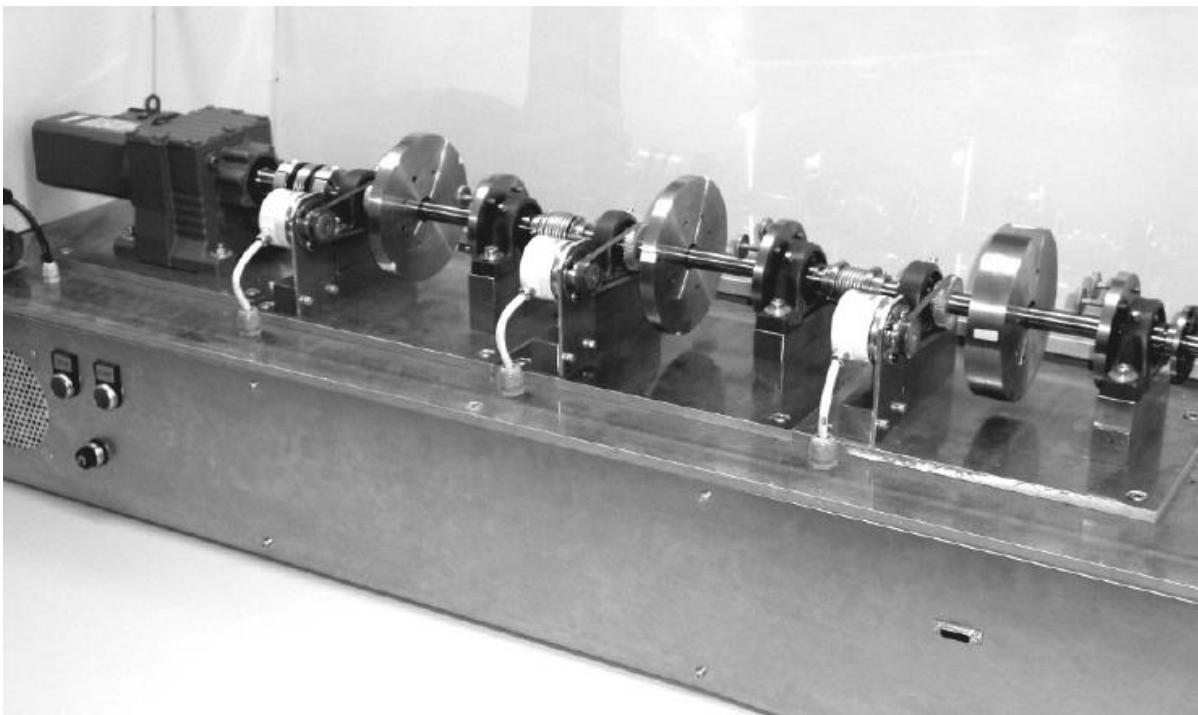
Wyznaczenie odpowiedzi częstotliwościowej

- Moduł błędu ma postać

$$|\Delta G_n(i\omega)| = \frac{|n(i\omega)|}{|u(i\omega)|}$$

- Moduł błędu (szumu) maleje wraz ze względnym wzrostem modułu wejścia nad szumem. Zatem gęstość amplitudowa musi być największa.
- Mogą to osiągnąć poprzez:
 - Wybór jak najwyższej amplitudy sygnału testowego (ograniczenia procesowe)
 - Dobór kształtu do określonych częstotliwości

Przykład: oscylator trzech mas



$$J_1 \ddot{\varphi}_1 = -d_1 \dot{\varphi}_1 - c_1 \varphi_1 + c_1 \varphi_2 + M_M$$

$$J_2 \ddot{\varphi}_2 = -d_2 \dot{\varphi}_2 + c_1 \varphi_1 - (c_1 + c_2) \varphi_2 + c_2 \varphi_3$$

$$J_3 \ddot{\varphi}_3 = -d_3 \dot{\varphi}_3 + c_2 \varphi_2 - c_2 \varphi_3 .$$

Isermann R (2005) Mechatronic Systems: Fundamentals. Springer, London

Przykład: oscylator trzech mas

- Równanie obiektu

$$J\ddot{\varphi}(t) + D\dot{\varphi}(t) + C\varphi = MM_M(t) \quad J = \begin{pmatrix} J_1 & 0 & 0 \\ 0 & J_2 & 0 \\ 0 & 0 & J_3 \end{pmatrix}, \quad D = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix}, \quad C = \begin{pmatrix} c_1 & -c_1 & 0 \\ -c_1 & (c_1 + c_2) & -c_2 \\ 0 & -c_2 & c_2 \end{pmatrix}$$

- Wektor stanu

$$x(t) = \begin{pmatrix} \varphi_1(t) \\ \varphi_2(t) \\ \varphi_3(t) \\ \dot{\varphi}_1(t) \\ \dot{\varphi}_2(t) \\ \dot{\varphi}_3(t) \end{pmatrix} \quad M = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

- Równania stanu

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -\frac{c_1}{J_1} & \frac{c_1}{J_1} & 0 & -\frac{d_1}{J_1} & 0 & 0 \\ \frac{c_1}{J_2} & -\frac{(c_1+c_2)}{J_2} & \frac{c_2}{J_2} & 0 & -\frac{d_2}{J_2} & 0 \\ 0 & \frac{c_2}{J_3} & -\frac{c_2}{J_3} & 0 & 0 & -\frac{d_3}{J_3} \end{pmatrix} \quad b = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{J_1} \\ 0 \\ 0 \end{pmatrix} \quad c^T = (0 \ 0 \ 1 \ 0 \ 0 \ 0)$$

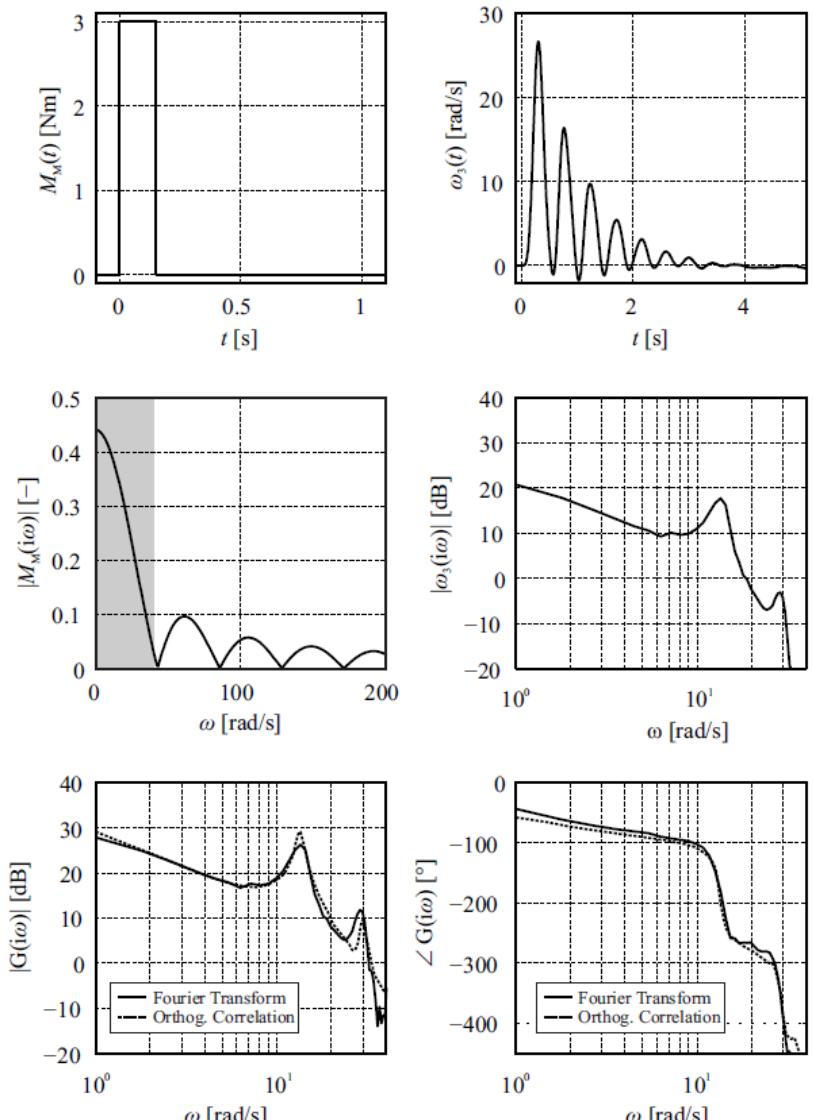
Przykład: oscylator trzech mas

- Transmitancja

$$\begin{aligned} G(s) &= \frac{\varphi_2(s)}{M_M(s)} \\ &= \frac{c_1}{J_1 J_2 s^4 + (J_2 d_1 + J_1 d_2)s^3 + (J_2 c_1 + d_1 d_2 + J_1 c_1)s^2 + (d_2 c_1 + d_1 c_1)s} \end{aligned}$$

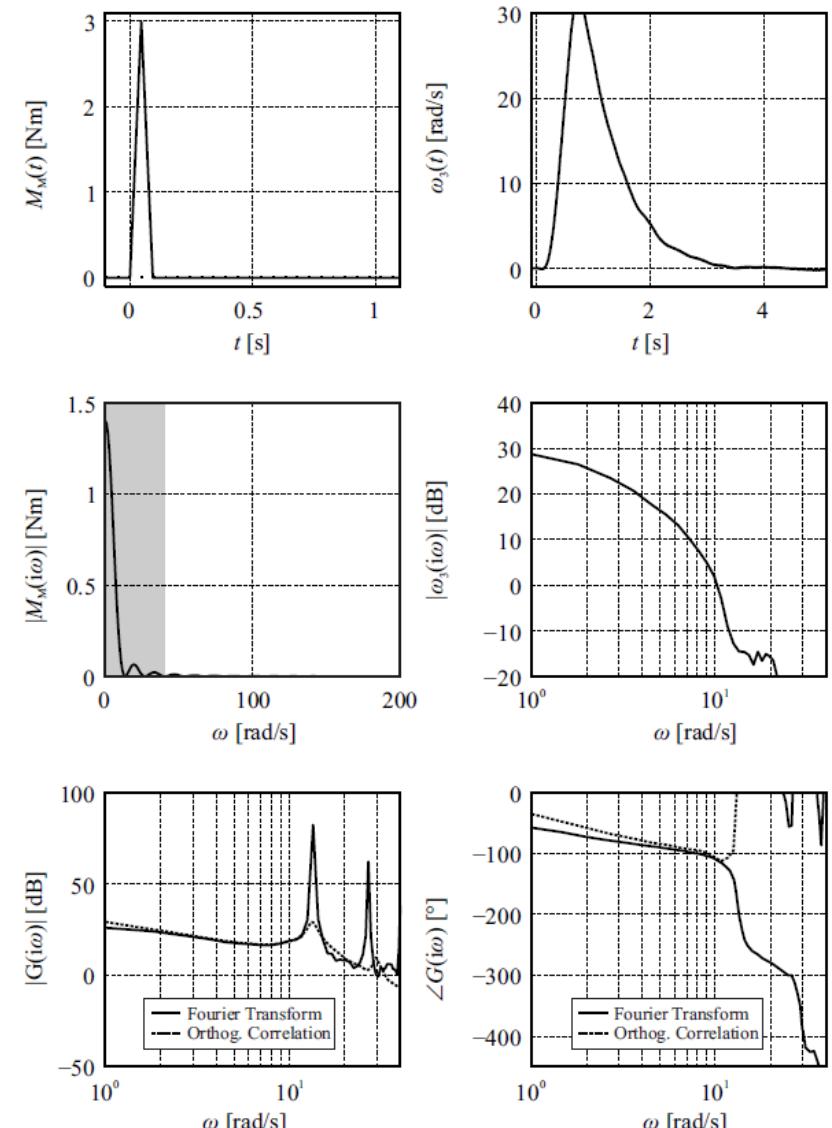
Przykład: oscylator trzech mas

- Sygnał identyfikacyjny – impuls prostokątny o okresie $T = 0.15$ s.
- Charakterystyka referencyjna pochodzi z metody korelacyjnej
- Dobre dopasowanie dla częstotliwości $\omega < 25$ [rad/s]



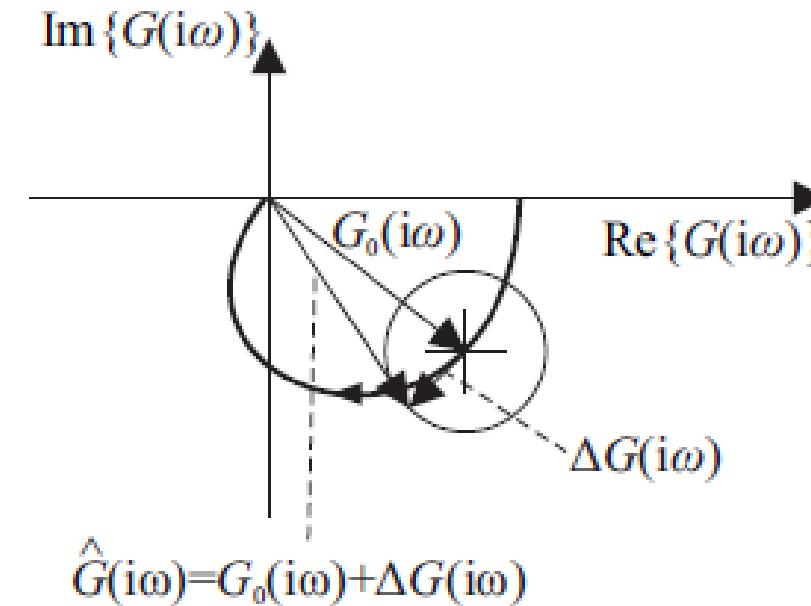
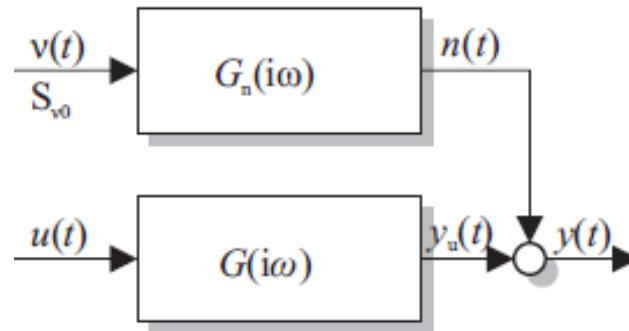
Przykład: oscylator trzech mas

- Sygnał identyfikacyjny – impuls trójkątny o okresie $T = 0.15$ s.
- Zero pokrywa się z częstotliwością maksimum charakterystyki
- Dobre dopasowanie dla mniejszego zakresu częstotliwości
 $\omega < 13.5$ [Hz]



Wpływy szumów pomiarowych

- Wykorzystanie jednorazowego pobudzenia jest tylko możliwe, gdy amplituda szumu jest znacznie mniejsza od impulsu oraz gdy średnia szumu jest równa zero.
- W przypadku znaczących niestacjonarnych szumów użycie powyższych metod może być w zasadzie niemożliwe.
- Zakładamy szum:



$$\Delta G_n(i\omega) = \frac{n_T(i\omega)}{u(i\omega)}$$

$$|\Delta G_n(i\omega)| = \frac{|n_T(i\omega)|}{|u(i\omega)|}$$

Wpływy szumów pomiarowych

- W wyniku przekształceń otrzymamy wniosek, że
 - błąd odpowiedzi częstotliwościowej jest odwrotnie proporcjonalny do stosunku sygnał-szum
 - błąd odpowiedzi częstotliwościowej jest odwrotnie proporcjonalny do pierwiastka z ilości wykorzystywanych odpowiedzi
- Tym samym w celu zmniejszenia wpływu szumu powinniśmy wyznaczyć wiele odpowiedzi na ten sam sygnał i wyznaczyć odpowiedź średnią:

$$\bar{y}(t) = \frac{1}{N} \sum_{k=0}^{N-1} y_k(t)$$

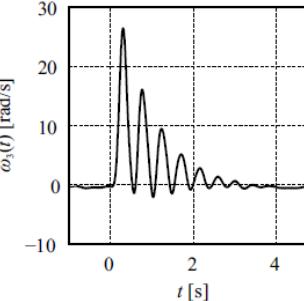
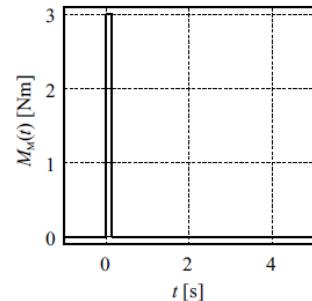
- Dla różnych pobudzeń powinniśmy wziąć średnią charakterystykę (uśredniając po części rzeczywistej i urojonej a nie po module i fazie)

$$\bar{G}(i\omega) = \frac{1}{N} \sum_{k=1}^N G_k(i\omega) = \frac{1}{N} \sum_{k=1}^N \operatorname{Re}\{G_k(i\omega)\} + i \frac{1}{N} \sum_{k=1}^N \operatorname{Im}\{G_k(i\omega)\}$$

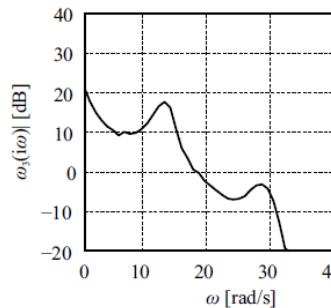
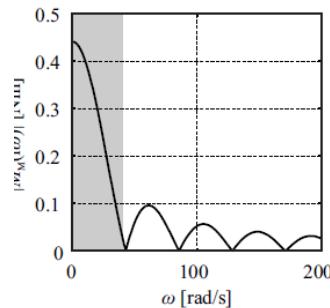
Przykład: oscylator trzech mas

- Stosunek szum do sygnału wynosi:

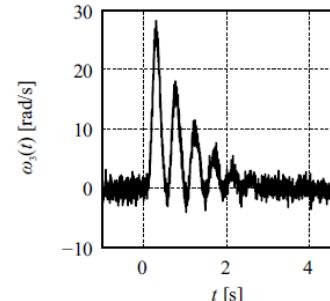
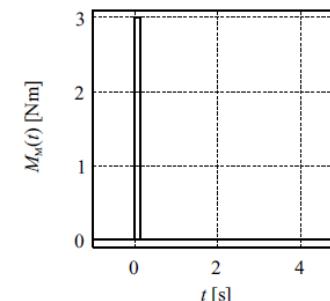
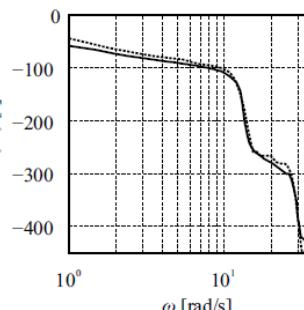
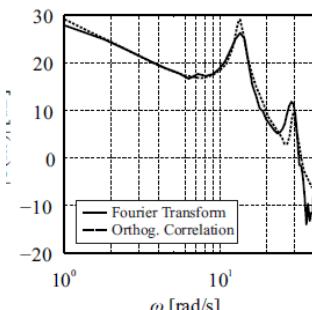
$$\eta = \frac{\sigma_n}{y_{\max}} \approx 4\% \triangleq 1 : 25$$



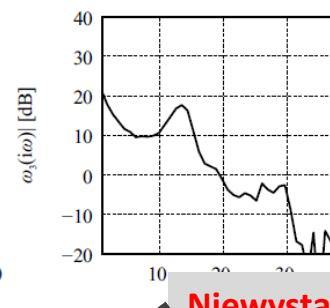
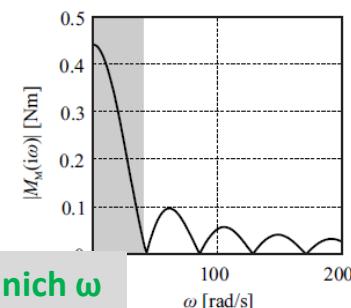
Bez szumu



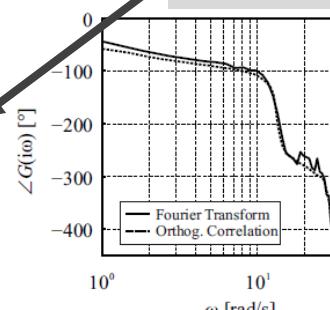
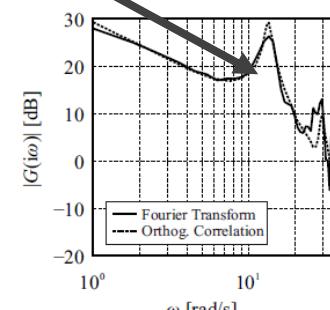
Całkiem nieźle w zakresie średnich ω



Z szumem

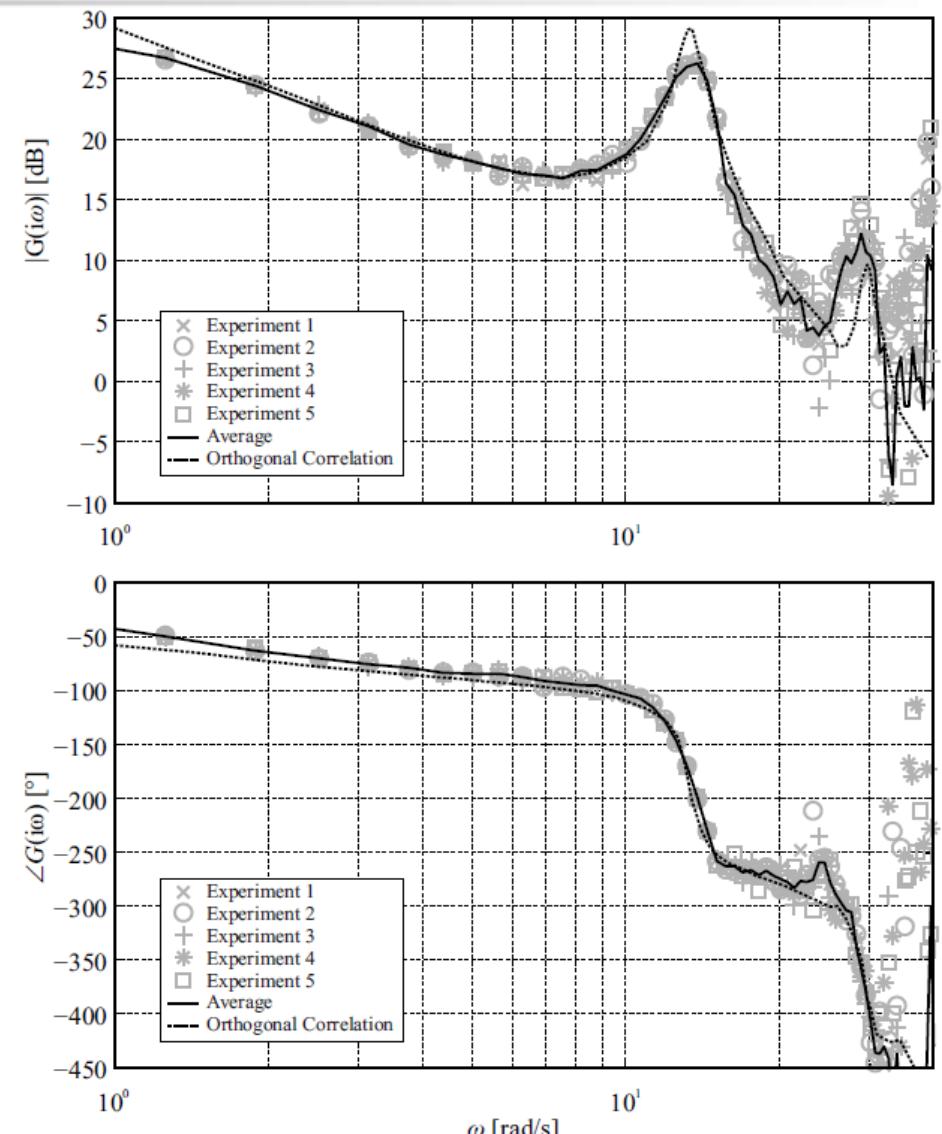


Niewystarczające dla wyższych ω



Przykład: oscylator trzech mas

- Uśrednienie wielu eksperymentów
 - Linia ciągła:
uśrednienie z pięciu eksperymentów
 - Linia przerywana:
rzeczywisty pomiar metoda korelacyjną
- Wariancja nie zanika dla $N \rightarrow \infty$



Podsumowanie

- Istotny jest rodzaj okienkowania oraz wybór sygnału testowego
- Określenie *preferowany sygnał* oznacza taki, który posiada najwyższą gęstość amplitudową w określonym zakresie częstotliwości
- Takie sygnały zapewniają najmniejszy błąd odtworzenia ch-ki częstotliwościowej w danym zakresie ω
- Dla niskich częstotliwości najlepszy jest skok jednostkowy a impuls prostokątny dla ω średnich
- Wymagana amplituda dla sygnału zaszumionego wynosi:

$$|u(i\omega)|_{\text{req}} = \frac{\sqrt{S_{nn}(i\omega)T_E}}{|G(i\omega)|\sigma_G(i\omega)\sqrt{N}}$$

- Niestety musimy znać proces

Podsumowanie

- Z punktu widzenia układów sterowania najistotniejsze są średnie częstotliwości
- Dobrze jest stosować różne sygnały:
 - Sekwencje skoków jednostkowych do odwzorowania niskich częstotliwości
 - Sekwencje impulsów prostokątnych dla niskich i średnich częstotliwości
 - 20%-30% czasu dla testów skokowych i 80%-70% czasu dla impulsów prostokątnych
 - Dobór długości impulsu prostokątnego zależy od najwyższej częstotliwości do odwzorowania
 $T = \pi/\omega_{max}$
 - Dla uwzględnienia potencjalnych nieliniowości sugeruje się stosowanie skoków w obu kierunkach
- W sytuacji wielokrotnego pobudzenia tym samym sygnałem można uśrednić odpowiedź w dziedzinie czasu i wtedy wyznaczyć charakterystykę
- W przypadku różnych sygnałów wyznaczamy oddzielnie charakterystyki dla poszczególnych eksperymentów i dopiero potem wyznaczamy uśrednione charakterystyki częstotliwościowe



MI

Metody Identyfikacji

wykład #5

1. *Wyznaczanie odpowiedzi częstotliwościowych dla okresowych sygnałów testowych*
2. *Analiza częstotliwościowa*

Odpowiedzi częstotliwościowe dla okresowych sygnałów testowych

- Możliwość wyznaczenia odpowiedzi dla pewnych dyskretnych punktów częstotliwości w określonym ich zakresie
 - Pobudzenie sygnałem sinusoidalnym o określonej częstotliwości
- Możliwość wykorzystania innych sygnałów:
 - Fala prostokątna
 - Fala trójkątna
 - Fala trapezowa
- Ręcznie lub komputerowo
 - Analiza korelacyjna
 - Analiza Fourierowska

Odpowiedzi częstotliwościowe dla okresowych sygnałów testowych

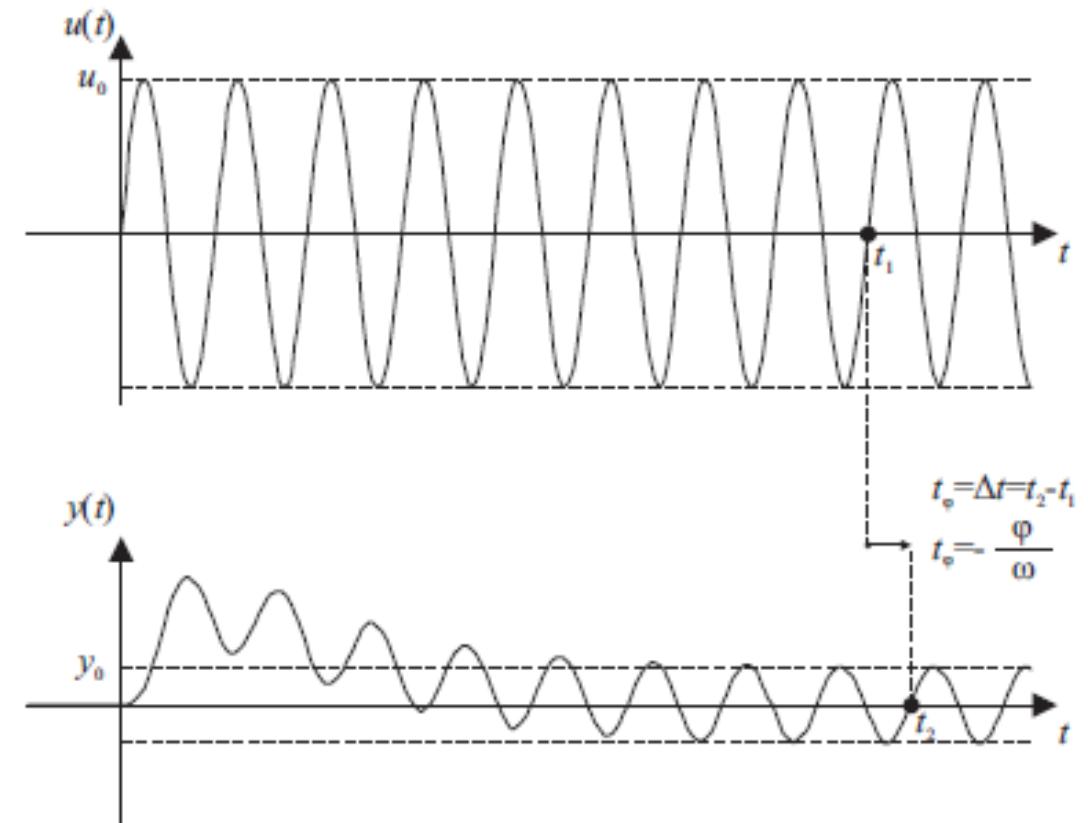
- Funkcja korelacji: możliwość wyznaczenia charakterystyk częstotliwościowych w obecności dużych zakłóceń i szumów
- Uwzględnienie charakterystyki urządzenia wykonawczego – **musi być liniowa**
 - Urządzenie wykonawcze z całkowaniem: użyć pozycjonera (układ regulacji) i sygnał przykładać jako wartość zadaną dla niego.
 - Urządzanie wykonawcze o stałych poziomach – sterowanie trzypoziomowe i fala prostokątna/trapezowa
 - Charakterystyki nieliniowe – sygnały dwupoziomowe – określone punkty pracy

Pobudzenie sinusoidalne

- Klasyczna metoda identyfikacji charakterystyk Bodego i Nyquista

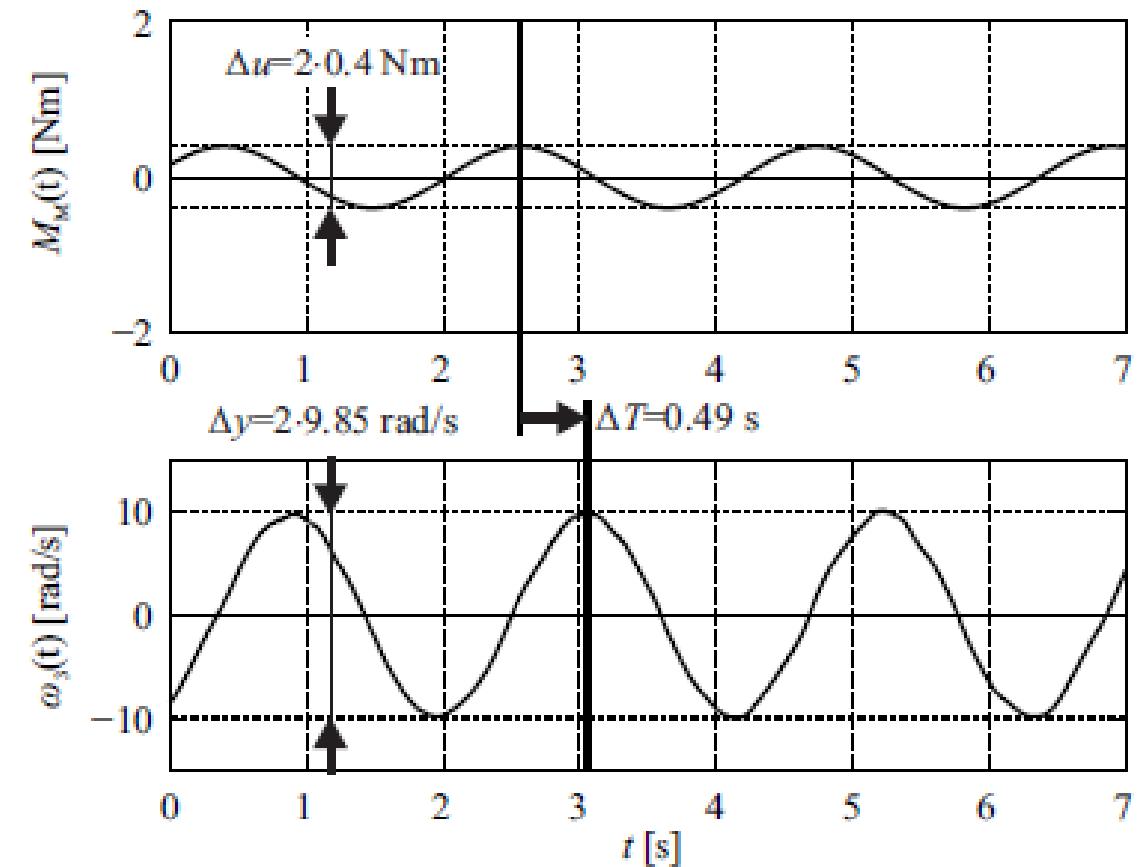
$$|G(i\omega_v)| = \frac{y_0(\omega_v)}{u_0(\omega_v)}$$

$$\angle G(i\omega_v) = -t_\varphi \omega_v ,$$



Pobudzenie sinusoidalne

- Przykład oscylatora trzech mas



Charakterystyki Bodego

- Uogólniona postać transmitancji

$$G(j\omega) = \frac{K \prod_{i=1}^Q (1 + j\omega\tau_i)}{(j\omega)^N \prod_{m=1}^M (1 + j\omega\tau_m) \prod_{k=1}^R \left(1 + \left(\frac{2\zeta_k}{\omega_{n_k}}\right)j\omega + \left(\frac{j\omega}{\omega_{n_k}}\right)^2\right)}$$

- Cztery podstawowe elementy:
 - Stałe wzmacnienie K
 - Biegun lub zero na osi ($j\omega$)
 - Zero lub biegun na osi rzeczywistej ($j\omega\tau+1$)
 - Zero lub biegun sprzężony $[1+(2\zeta/\omega_n)j\omega+(\omega_n/j\omega)^2]$
- Charakterystyki Bodego dla tychże elementów

Pobudzenie falą prostokątną (trapezową)

- Fala prostokątna może zostać opisana za pomocą szeregu Fouriera $\omega_o = 2\pi/T$

$$u(t) = \frac{4}{\pi} u_0 \left(\sin \omega_0 t + \frac{1}{3} \sin 3\omega_0 t + \frac{1}{5} \sin 5\omega_0 t + \dots \right)$$

- Zaś odpowiedź ma postać

$$\begin{aligned} y(t) = & \frac{4}{\pi} u_0 \left(|G(i\omega_0)| \sin(\omega_0 t + \varphi(\omega_0)) \right. \\ & + \frac{1}{3} |G(i3\omega_0)| \sin(3\omega_0 t + \varphi(3\omega_0)) \\ & \left. + \frac{1}{5} |G(i5\omega_0)| \sin(5\omega_0 t + \varphi(5\omega_0)) + \dots \right) \end{aligned}$$

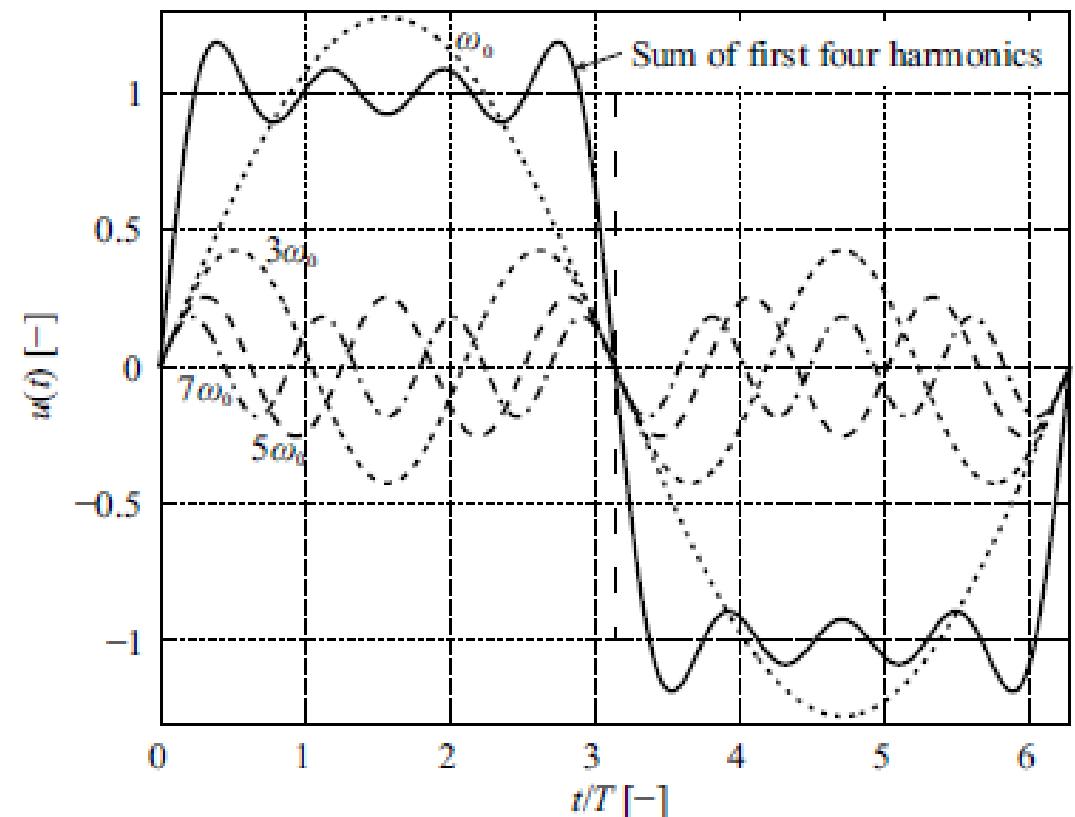
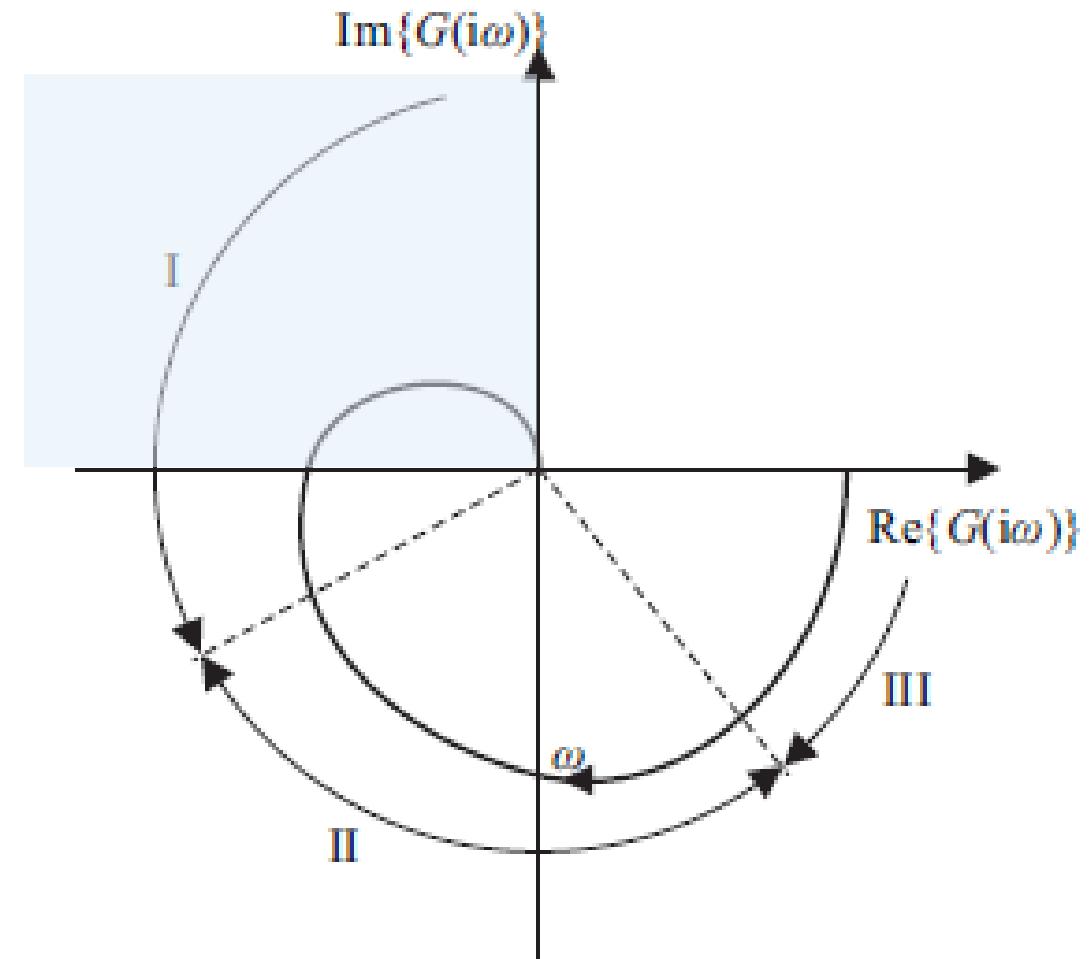


Fig. 5.3. Harmonic decomposition of a rectangular wave

Pobudzenie falą prostokątną (trapezową)

- Zaczynamy od zakresu wyższych częstotliwości
- Wyższe harmoniczne $n \geq 3$ są tłumione tak mocno, że odpowiedź odpowiada prawie że czystej sinusoidzie
- Łatwo można odczytać moduł i przesunięcie fazowe odpowiedzi
- Odczytujemy część I charakterystyki Nyquista



Pobudzenie falą prostokątną (trapezową)

- Dla średnich częstotliwości amplituda drugiej harmonicznej dla $3\omega_0$ osiąga wielkość, która nie może być pominięta

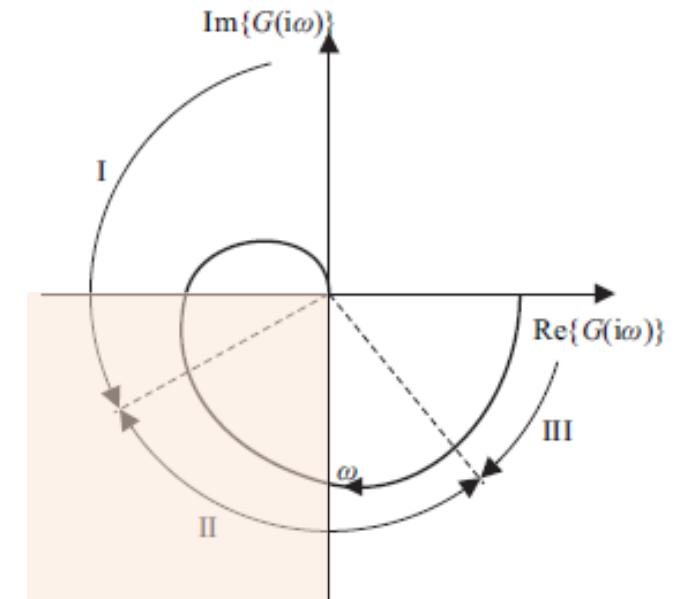
$$y(t) \approx \frac{4}{\pi} u_0 \left(|G(i\omega_0)| \sin(\omega_0 t + \varphi(\omega_0)) + \frac{1}{3} |G(i3\omega_0)| \sin(3\omega_0 t + \varphi(3\omega_0)) \right)$$

- Trzecia harmoniczna $5\omega_0$ może zostać pominięta tak jak i wyższe. Zatem można otrzymać odpowiedź dla danej częstotliwości odejmując odpowiedź dla drugiej składowej harmonicznej

$$y_{3\omega_0}(t) = \frac{4}{\pi} \frac{1}{3} u_0 |G(i3\omega_0)| \sin(3\omega_0 t + \varphi(3\omega_0))$$

od mierzonego wyjścia $y(t)$.

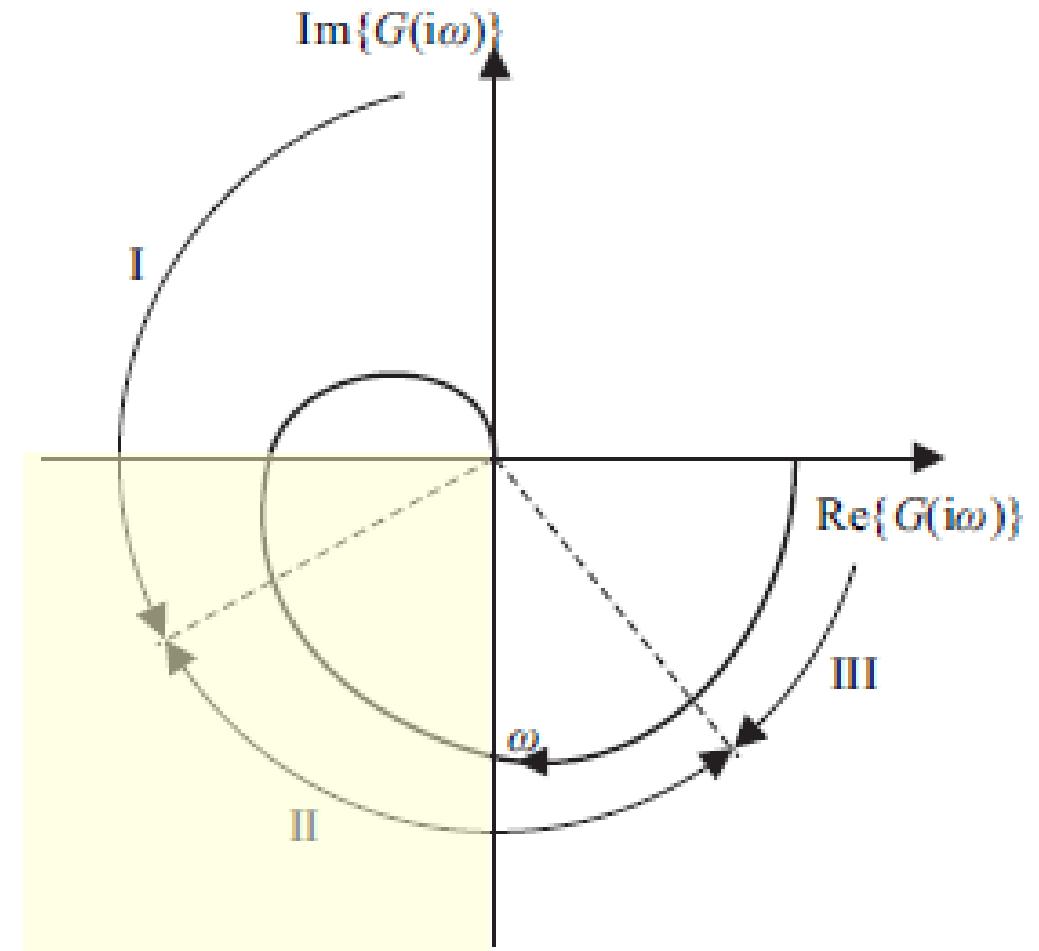
- A odpowiedź $y_{3\omega_0}$ jest znana z poprzedniego wywodu.
- Tym samym otrzymujemy strefę II.



Pobudzenie falą prostokątną (trapezową)

- Dla niskich częstotliwości odpowiedź otrzymujemy odejmując tak wiele harmonicznych jak trzeba.

$$\begin{aligned} & \frac{4}{\pi} u_0 |G(i\omega_0)| \sin(\omega_0 t + \varphi(\omega_0)) \\ &= y - u_0 \frac{1}{3} |G(i3\omega_0)| \sin(3\omega_0 t + \varphi(3\omega_0)) \\ & - u_0 \frac{1}{5} |G(i5\omega_0)| \sin(5\omega_0 t + \varphi(5\omega_0)) - \dots \end{aligned}$$



Pobudzenie falą prostokątną (trapezową)

- Metoda głównie stosowana dla wyższych zakresów częstotliwości
- Podsumowanie:
 - Łatwiej zrealizować fale prostokątną niż sinusoidalną
 - Charakterystyka urządzenia wykonawczego nie musi być liniowa
 - Dla danej amplitudy u_o fala prostokątna ma najwyższą amplitudę podstawowej harmonicznej niż dla jakichkolwiek innych pobudzeń okresowych. Zatem najlepszy stosunek do zakłóceń.
 - Ze względu na charakter przekształcenia Fouriera czas przełączenia od $+u_o$ do $-u_o$ powinien być nie większy niż:

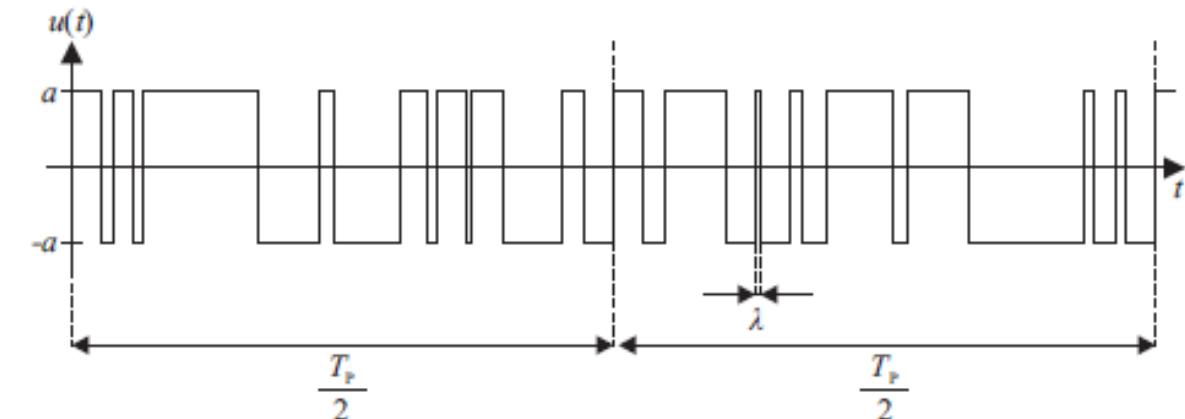
$$T_1^* < \frac{1.1}{\omega_{\max}} \text{ resp. } T_1^* < \frac{0.5}{\omega_{\max}}$$

- Gdzie ω_{\max} to najwyższa częstotliwość w interesującym zakresie

Pobudzenie z wieloma harmonicznymi

- Binarny sygnał o wielu częstotliwościach
- Cel: maksymalizacja amplitud dla poszczególnych składowych częstotliwościowych przy minimalizacji amplitudy właściwego pobudzenia

$$\left. \begin{aligned} |G(i\omega_v)| &= \frac{1}{u_{0v}} \sqrt{a_{yyv}^2 + b_{yyv}^2} \\ \varphi(\omega_v) &= \arctan \frac{a_{yyv}}{b_{yyv}} \end{aligned} \right\}$$



Sześć (6) harmonicznych

$$u_o = 0.585a$$

$$N = 256$$

$$\omega_o, 2\omega_o, 4\omega_o, 8\omega_o, 16\omega_o, 32\omega_o$$

Sygnały z ciągłą zmianą częstotliwości

Equation

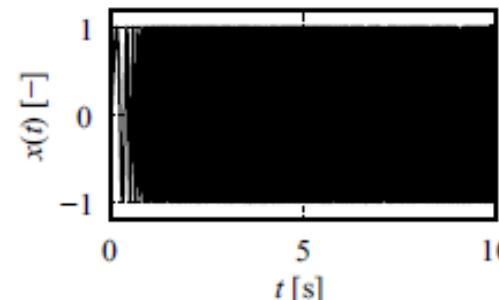
$$x(t) = \sin(2\pi f(t) t)$$

with

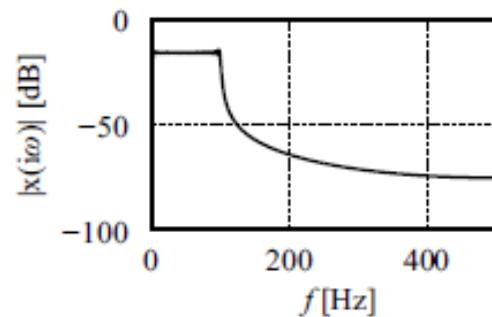
$$f(t) = f_0 + \frac{f_1 - f_0}{2T} t$$

Linear Sweep

Time Response



Fourier Transform



Equation

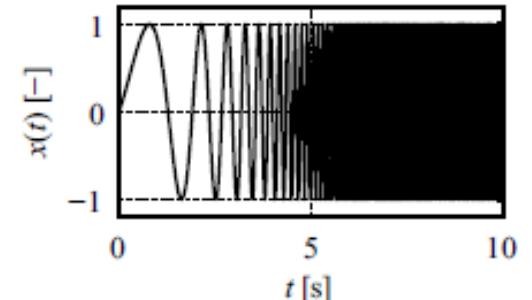
$$x(t) = \sin(2\pi f(t) t)$$

with

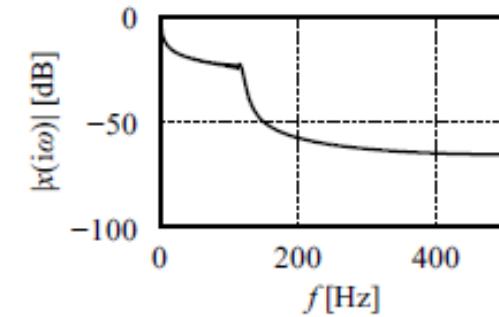
$$f(t) = f_0 \left(\frac{f_1}{f_0}\right)^{\frac{t}{T}}$$

Exponential Sweep

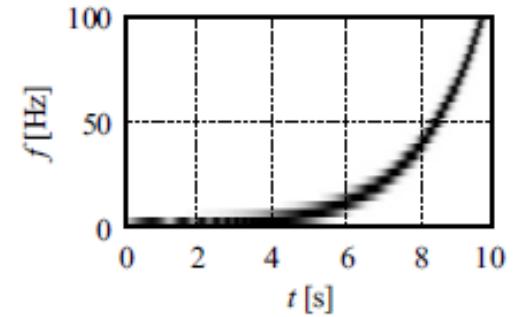
Time Response



Fourier Transform



Spectrogram



Stosowane do analizy w teorii obwodów i identyfikacji sieci (ogólnie telekomunikacja)

Metoda koreacyjna dla charakterystyk częstotliwościowych

- Funkcja autokorelacji

$$R_{uu}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} u(t)u(t + \tau)dt$$

- Funkcja korelacji wzajemnej

$$\begin{aligned} R_{uy}(\tau) &= E\{u(t)y(t + \tau)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} u(t)y(t + \tau)dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} u(t - \tau)y(t)dt . \end{aligned}$$

- Obie funkcje są ze sobą powiązane

$$R_{uy}(\tau) = \int_0^{\infty} g(t')R_{uu}(\tau - t')dt'$$

- Ważne nie tylko dla sygnałów stochastycznych ale i okresowych

Metoda koreacyjna dla charakterystyk częstotliwościowych

- Możliwość wyznaczenia odpowiedzi impulsowej (poprzez znalezienie funkcji autokorelacji oraz korelacji wzajemnej)
- Zaś na podstawie odpowiedzi impulsowej (transformata Fouriera) można znaleźć charakterystyki częstotliwościowe
- Ze względu na poniższe cechy **można bezpośrednio wyznaczyć moduł i fazę!**
- Dla pobudzenia sinusoidalnego

$$u(t) = u_o \sin \omega_o t, \text{ o częstotliwości } \omega_o = 2\pi/T_p$$

- Funkcja autokorelacji przyjmuje postać

$$R_{uu}(\tau) = \frac{2u_0^2}{T_0} \int_0^{\frac{T_0}{2}} \sin(\omega_0 t + \alpha) \sin(\omega_0(t + \tau) + \alpha) dt = \frac{u_0^2}{2} \cos \omega_0 \tau$$



Metoda korelacyjna dla charakterystyk częstotliwościowych

- Funkcja korelacji wzajemnej sygnału testowego i odpowiedzi na sygnał testowy

$$y(t) = u_0 |G(i\omega_0)| \sin(\omega_0 t - \varphi(\omega_0))$$

- prowadzi do postaci

$$\begin{aligned} R_{uy}(\tau) &= |G(i\omega_0)| \frac{2u_0^2}{T_p} \int_0^{\frac{T_p}{2}} \sin \omega_0(t-\tau) \sin(\omega_0 t - \varphi(\omega_0)) dt \\ &= |G(i\omega_0)| \frac{u_0^2}{2} \cos(\omega_0 \tau - \varphi(\omega_0)). \end{aligned}$$

- Ze względu na okresowość całkowanie można ograniczyć do połowy okresu.
- Ostatecznie otrzymujemy:

$$R_{uy}(\tau) = |G(i\omega_0)| R_{uu}\left(\tau - \frac{\varphi(\omega_0)}{\omega_0}\right)$$

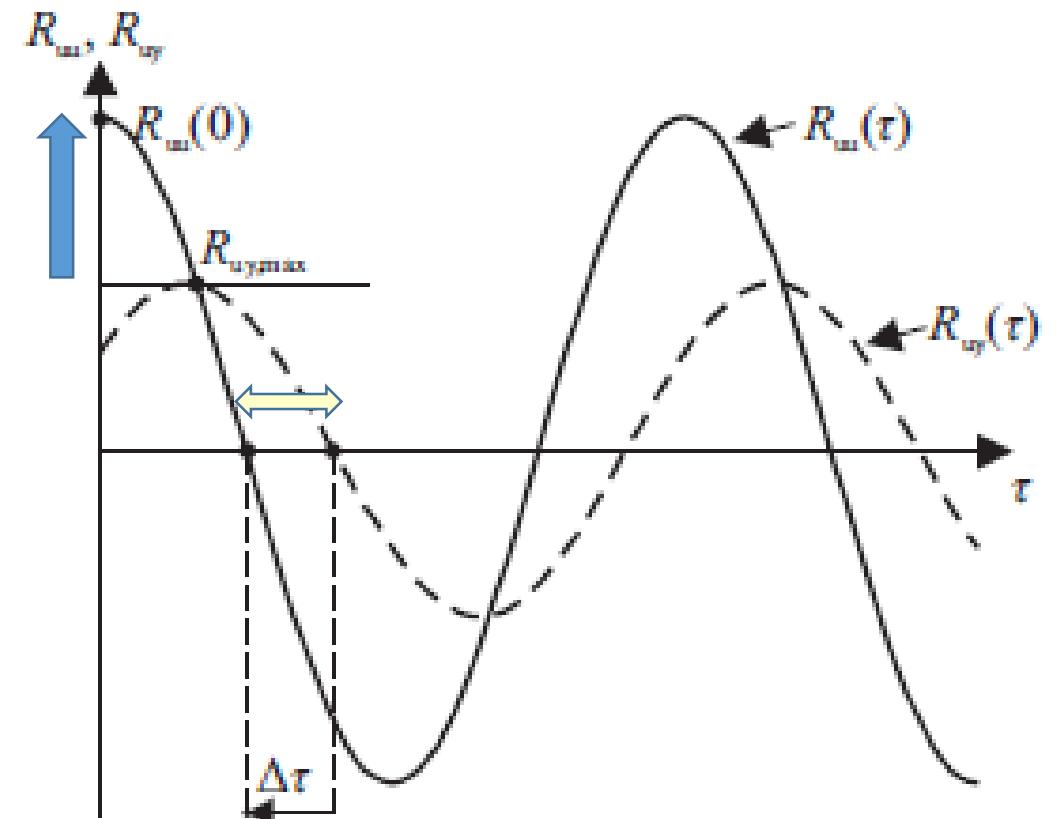
Metoda koreacyjna dla charakterystyk częstotliwościowych

- Rysując obie korelacje w funkcji czasu otrzymujemy:
 - Moduł charakterystyki to stosunek korelacji wzajemnej w punkcie τ do funkcji autokorelacji w momencie $\tau - \varphi(\omega_o)/\omega_o$

$$|G(i\omega_0)| = \frac{R_{uy}(\tau)}{R_{uu}\left(\tau - \frac{\varphi(\omega_0)}{\omega_0}\right)} = \frac{R_{uy,\max}}{R_{uu}(0)} = \frac{R_{uy}\left(\frac{\varphi(\omega_0)}{\omega_0}\right)}{R_{uu}(0)}$$

- A faza jest określana z opóźnienia $\Delta\tau$

$$\varphi(\omega_o) = -\omega_o \Delta\tau$$



Metoda koreacyjna dla charakterystyk częstotliwościowych

- Metoda nie jest ograniczona dla sygnałów sinusoidalnych
- Mogą być sygnały o wielu harmonicznych dopóki korelacje są sinusoidalne
- W obecności zakłóceń należy rozpatrywać dłuższe przebiegi – wiele okresów
 - Błąd zanika jeśli zakłócenie stochastyczne nie jest skorelowane z wejściem
 - Dotyczy to również przebiegów okresowych, tyle że muszą mieć częstotliwość różną od ω_o

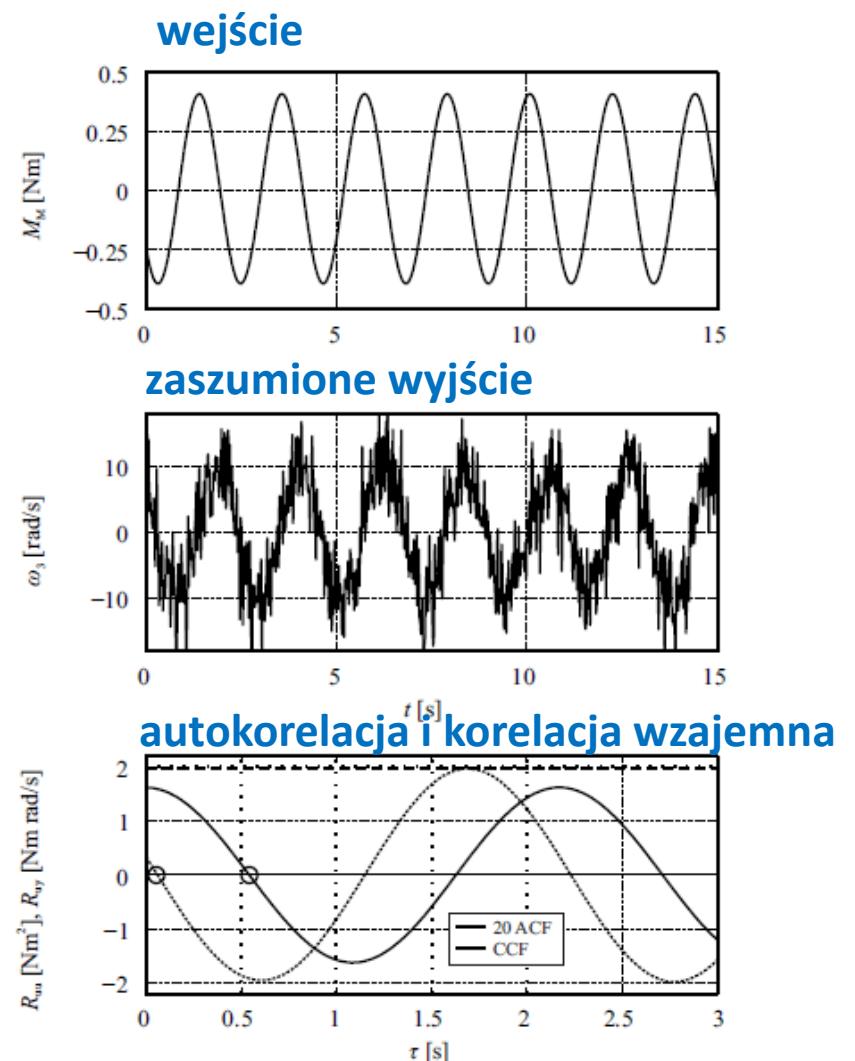
Przykład oscylatora trzech mas

- Otrzymujemy moduł

$$|G(i\omega_0)|_{\omega=2.8947 \text{ rad/s}} = \frac{R_{uy,\max}}{R_{uu}(0)} = \frac{1.99 \text{ Nm} \frac{\text{rad}}{\text{s}}}{0.081 \text{ Nm}^2} = 24.49 \frac{\text{rad}}{\text{s Nm}}$$

- oraz fazę

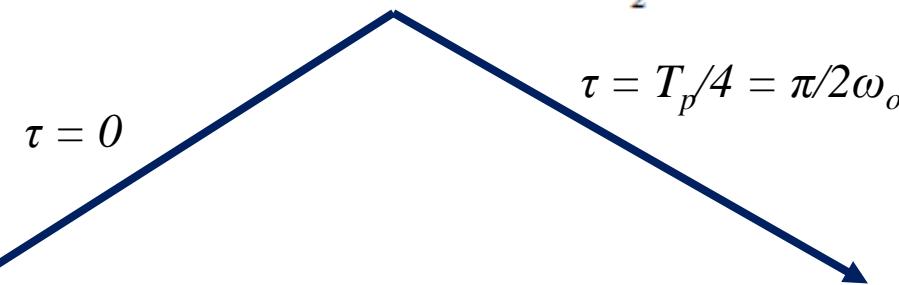
$$\begin{aligned} \varphi(i\omega)|_{\omega=2.8947 \text{ rad/s}} &= -\Delta \tau \omega = (0.542 \text{ s} - 0.053 \text{ s}) 2.8947 \frac{\text{rad}}{\text{s}} \\ &= -1.41 \text{ rad} = -81.1^\circ. \end{aligned}$$



Korelacja ortogonalna

- Można wyznaczyć część rzeczywistą i urojoną charakterystyki z funkcji korelacji wzajemnej

$$|G(i\omega_0)| \cos(\omega_0 \tau - \varphi(\omega_0)) = \frac{R_{uy}(\tau)}{\frac{u_0^2}{2}}$$



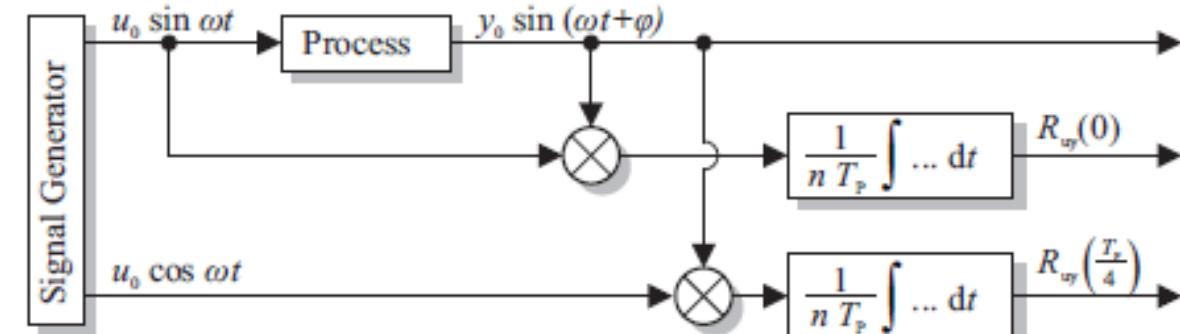
$$\operatorname{Re}\{G(i\omega_0)\} = |G(i\omega_0)| \cos(\varphi(\omega_0)) = \frac{R_{uy}(0)}{\frac{u_0^2}{2}}$$

$$R_{uy}(0) = \frac{u_0^2}{2} \operatorname{Re}\{G(i\omega_0)\} = \frac{u_0}{n T_p} \int_0^{n T_p} y(t) \sin \omega_0 t \, dt$$

$$\operatorname{Im}\{G(i\omega_0)\} = |G(i\omega_0)| \sin(\varphi(\omega_0)) = \frac{R_{uy}\left(\frac{\pi}{2\omega_0}\right)}{\frac{u_0^2}{2}}$$

$$R_{uy}\left(\frac{T_p}{4}\right) = \frac{u_0^2}{2} \operatorname{Im}\{G(i\omega_0)\} = -\frac{u_0}{n T_p} \int_0^{n T_p} y(t) \cos \omega_0 t \, dt$$

- Zatem wystarcza wyznaczyć funkcję korelacji wzajemnej jedynie w dwu punktach



Korelacja ortogonalna

- Dzięki ortogonalności funkcji trygonometrycznych otrzymujemy podstawowe zależności:

$$\begin{aligned} R_{uy}(0) &= \frac{1}{nT_p} u_0 y_0 \left(\int_0^{nT_p} \sin^2 \omega_0 t \cos \varphi dt + \underbrace{\int_0^{nT_p} \sin \omega_0 t \cos \omega_0 t \sin \varphi dt}_{=0} \right) \\ &= \frac{y_0}{u_0} \frac{u_0^2}{2} \cos \varphi = |G(i\omega_0)| \cos \varphi \frac{u_0^2}{2} = \text{Re}\{G(i\omega_0)\} \frac{u_0^2}{2}. \end{aligned}$$

$$\begin{aligned} R_{uy}\left(\frac{T_p}{4}\right) &= \frac{1}{nT_p} \int_0^{nT_p} u_0 \cos \omega_0 t y_0 \sin(\omega_0 t + \varphi) dt \\ &= \text{Im}\{G(i\omega_0)\} \frac{u_0^2}{2}. \end{aligned}$$

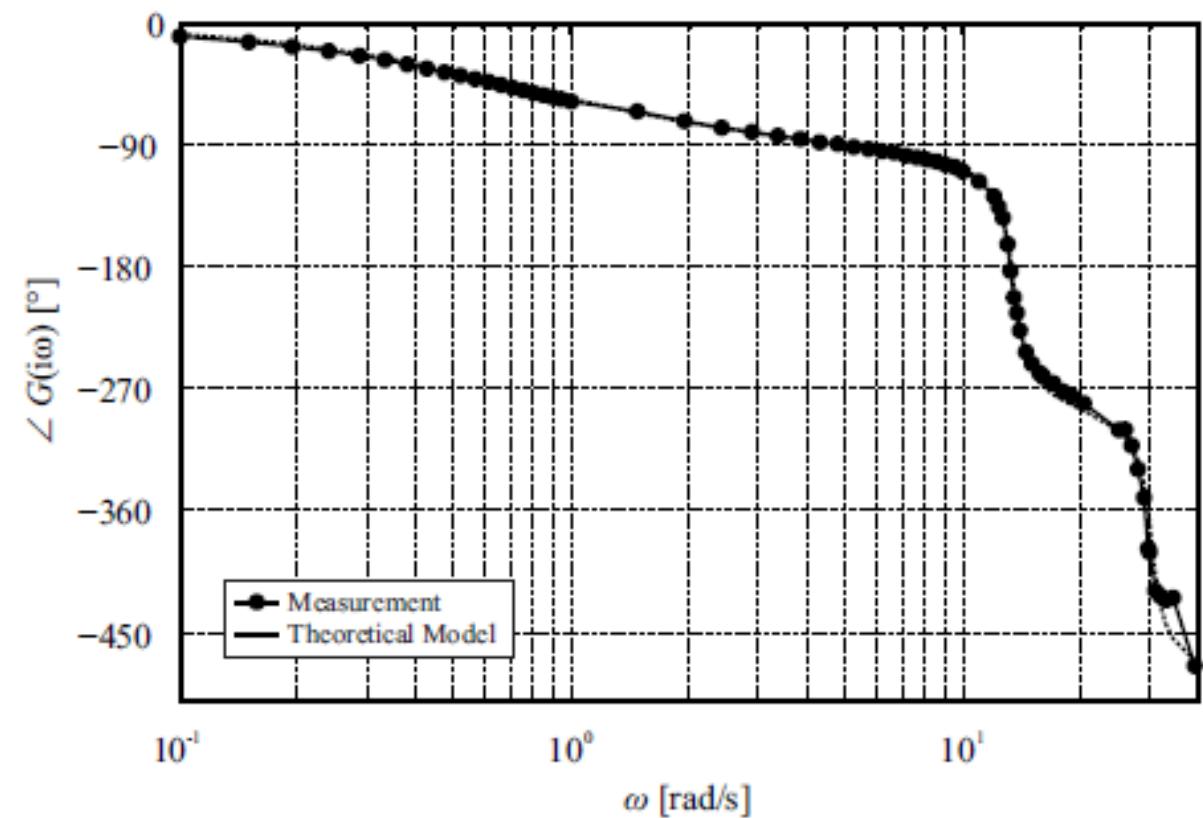
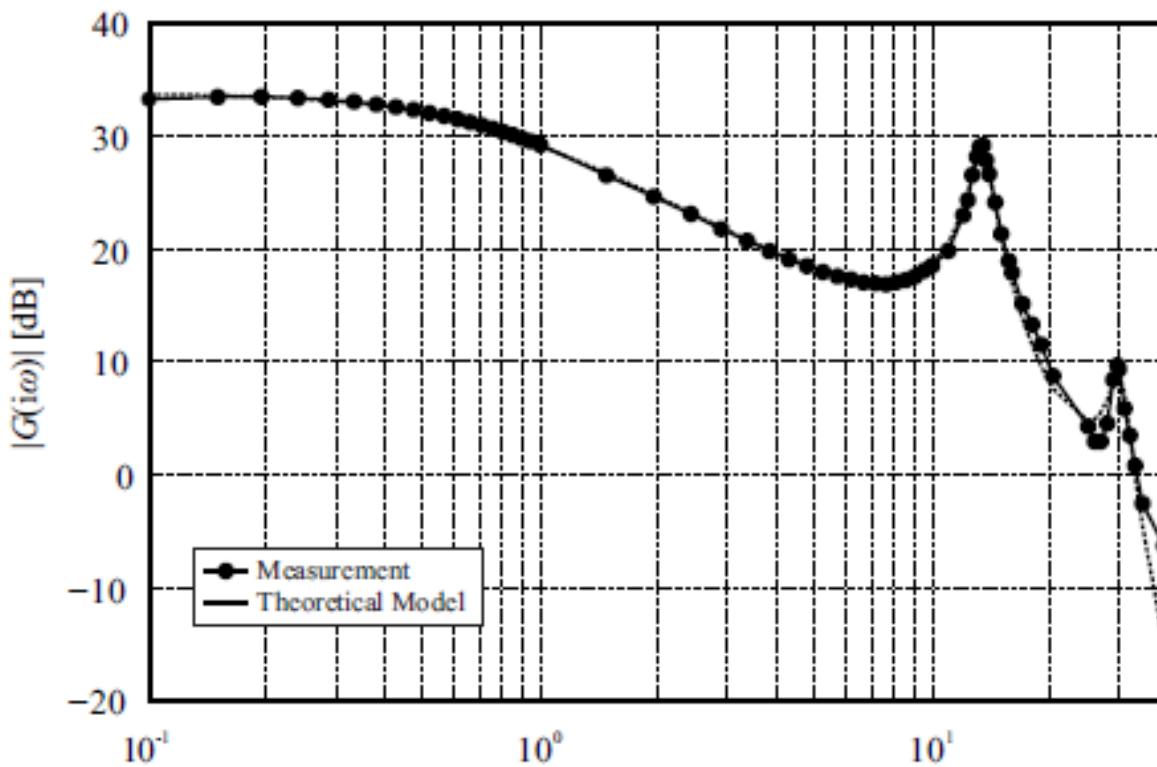
- Amplituda i faza otrzymana jest w postaci

$$|G(i\omega_0)| = \sqrt{\text{Re}^2\{G(i\omega_0)\} + \text{Im}^2\{G(i\omega_0)\}}$$

$$\varphi(\omega_0) = \arctan \frac{\text{Im}\{G(i\omega_0)\}}{\text{Re}\{G(i\omega_0)\}}.$$

- Metoda klasyczna – nie jest zakłócana częstotliwościami innymi od ω_o

Przykład: oscylator trzech mas

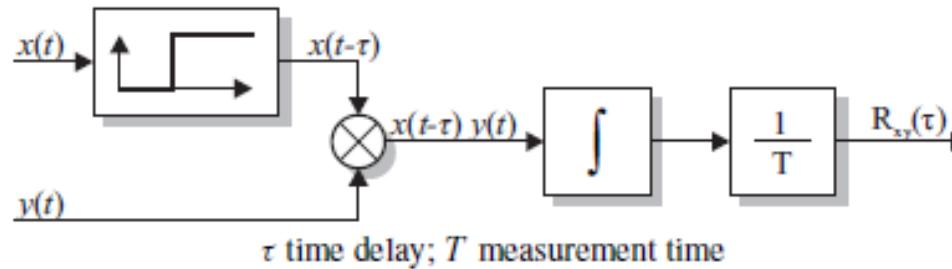


Metoda korelacyjna

- Praktyczne zastosowanie w celu ograniczenia czasu doświadczenia:
 - Niskie i średnie częstotliwości → analiza Fourierowska
 - Wysokie częstotliwości → metoda korelacyjna
- Co dalej?

Pobudzanie sygnałami zawierającymi wiele harmonicznych!!!

Estymacja funkcji korelacji wzajemnej



$$R_{xy}(\tau) = E\{x(t)y(t + \tau)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)y(t + \tau)dt$$

$$= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t - \tau)y(t)dt$$

$$R_{yx}(\tau) = E\{y(t)x(t + \tau)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} y(t)x(t + \tau)dt$$

$$= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} y(t - \tau)x(t)dt$$

$$R_{xy}(\tau) = -R_{yx}(\tau)$$

Wpływ długości okresu T eksperymentu jest istotny!!!

Estymacja funkcji korelacji wzajemnej

- Wpływy zakłóceń zanikają wraz z przyrostem długości okresu T
- Wariancja estymaty korelacji wzajemnej zanika odwrotnie proporcjonalnie do czasu T
- Estymata korelacji wzajemnej jest nieobciążona jeśli zakłócenie addytywne na wejściu i wyjściu jest odpowiednio niezależne od wejścia i wyjścia
 - Wariancja estymaty jest wtedy równa sumie wariancji obu zakłóceń
 - **Oba zakłócenia muszą być od siebie niezależne**

Estymacja funkcji autokorelacji

- Estymata jest nieobciążona
- Jeśli sygnał jest zakłócony

$$x(t) = x_o(t) + n(t)$$

- otrzymujemy

$$R_{xx}(\tau) = R_{xoxo}(\tau) + R_{nn}(\tau)$$

Estymaty funkcji korelacji wzajemnej i autokorelacji

- Obie funkcje korelacji i korelacji wzajemnej dla obiektów liniowych o odpowiedzi impulsowej $g(t)$ są estymowane w sensie minimalnokwadratowym przy następujących warunkach:
 - Sygnały wejściowy i wyjściowy jest **stacjonarny**
 - Wartość oczekiwana wejścia **wynosi zero**
 - Zakłócenie jest **stacjonarne i nieskorelowane** z wejściem
 - Analogicznie jak niezależnie zakłócone addytywnie jest wejście i wyjście, ale oba zakłócenia muszą być **nieskorelowane wzajemnie**



MI

Metody Identyfikacji

wykład #5-a

1. *Metody częstotliwościowe – pobudzanie sygnałem sinusoidalnym*
2. *Bode, Nyquist, Nichols*
3. *Inne sygnały periodyczne*

Odpowiedzi częstotliwościowe dla okresowych sygnałów testowych

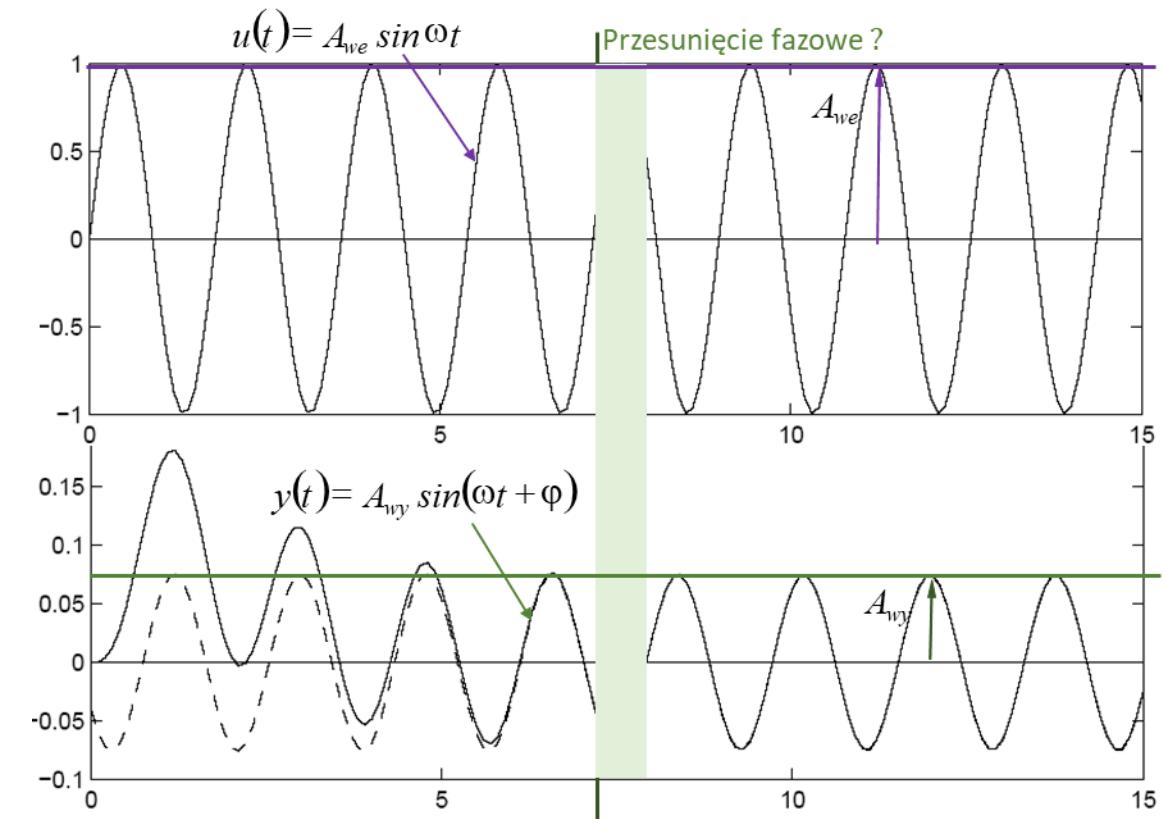
- Możliwość wyznaczenia odpowiedzi dla pewnych dyskretnych punktów częstotliwości w określonym ich zakresie
 - Pobudzenie sygnałem sinusoidalnym o określonej częstotliwości
- Możliwość wykorzystania innych sygnałów:
 - Fala prostokątna
 - Fala trójkątna
 - Fala trapezowa
- Ręcznie lub komputerowo
 - Analiza korelacyjna
 - Analiza Fourierowska

Odpowiedzi częstotliwościowe dla okresowych sygnałów testowych

- Funkcja korelacji: możliwość wyznaczenia charakterystyk częstotliwościowych w obecności dużych zakłóceń i szumów
- Uwzględnienie charakterystyki urządzenia wykonawczego – **musi być liniowa**
 - Urządzenie wykonawcze z całkowaniem: użyć pozycjonera (układ regulacji) i sygnał przykładać jako wartość zadaną dla niego.
 - Urządzanie wykonawcze o stałych poziomach – sterowanie trzypoziomowe i fala prostokątna/trapezowa
 - Charakterystyki nieliniowe – sygnały dwupoziomowe – określone punkty pracy

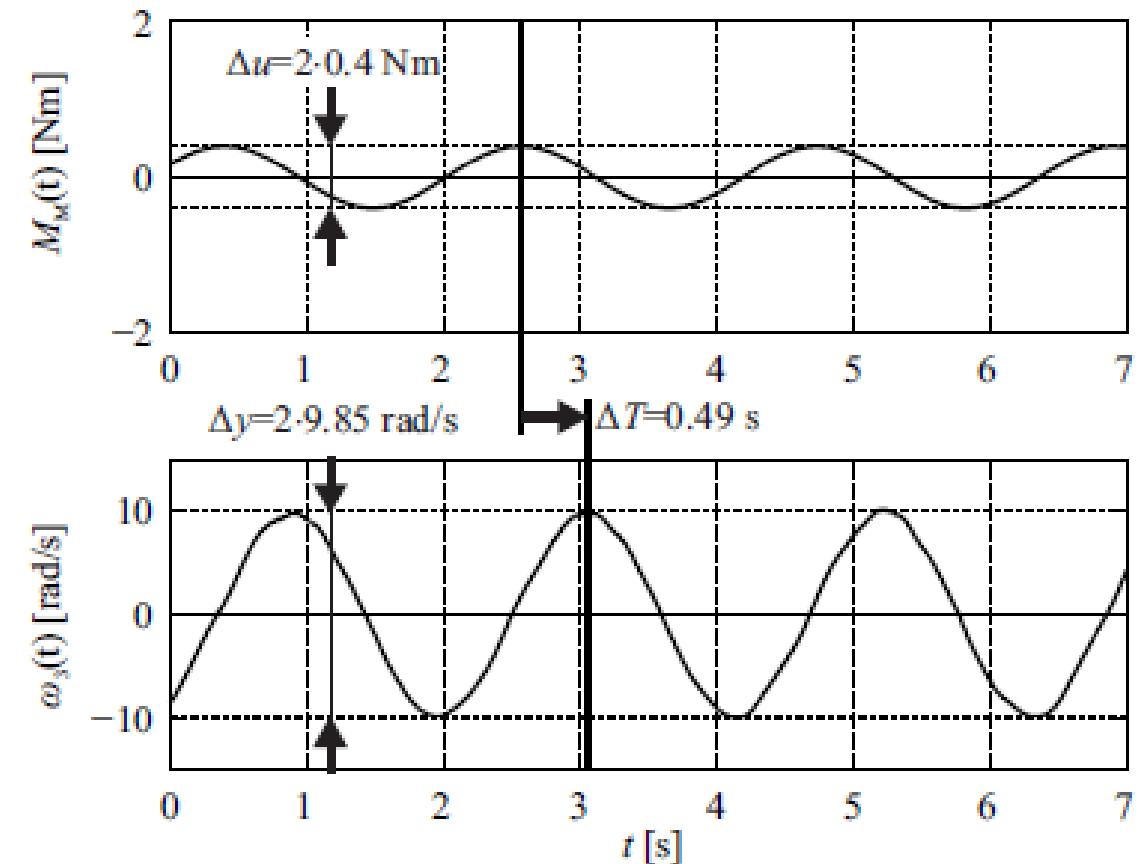
Pobudzenie sinusoidalne

- Klasyczna metoda identyfikacji charakterystyk Bodego i Nyquista
 - Moduł → stosunek amplitudy sinusoidy na wyjściu do amplitudy sinusoidy na wejściu
 - Faza → przesunięcie fazy pomiędzy sinusoidą na wyjściu do sinusoidy na wejściu
 - $\angle G(j\omega) = \angle \frac{Y(j\omega)}{X(j\omega)}$



Pobudzenie sinusoidalne

- Przykład oscylatora trzech mas



Charakterystyki Bodego

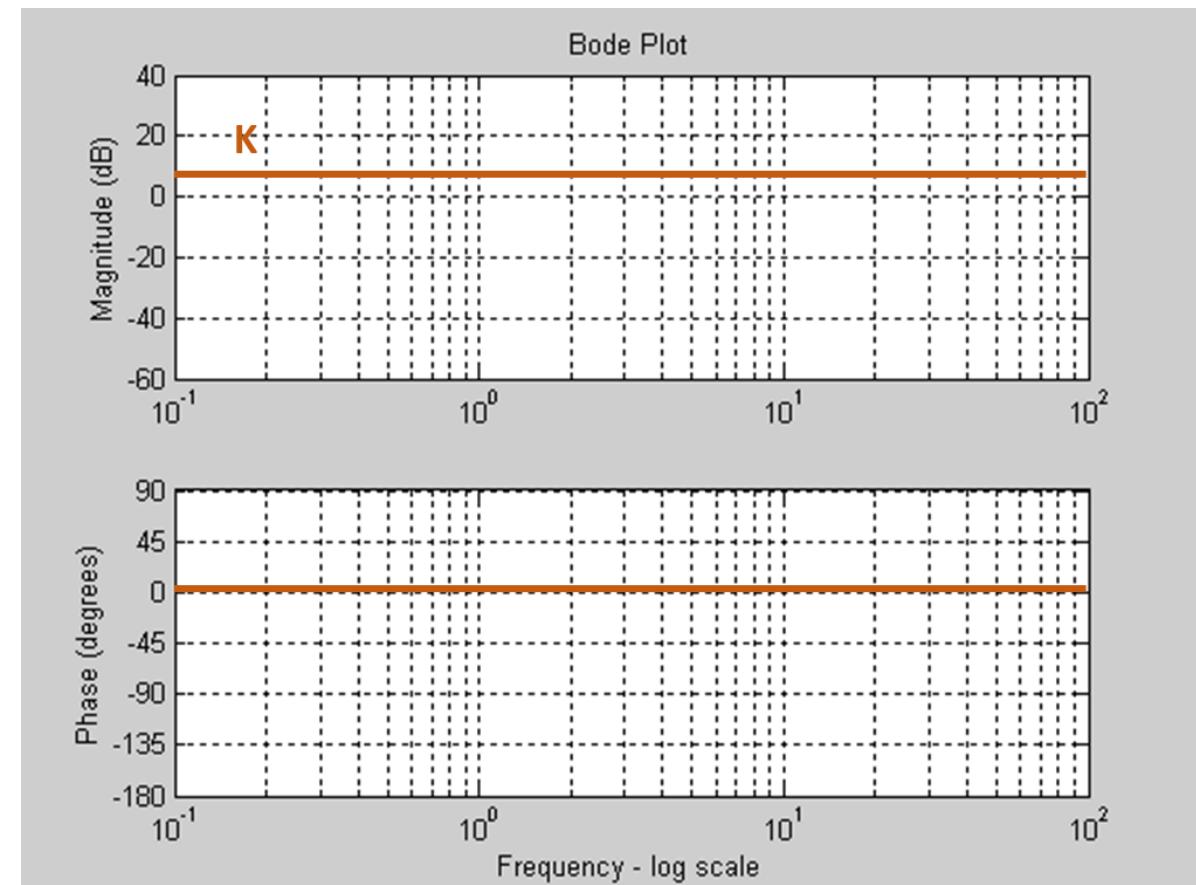
- Uogólniona postać transmitancji
- Wpisz tutaj równanie.

$$G(j\omega) = \frac{K}{(j\omega)^n} \frac{\prod_i^P (1 + j\omega\tau_i) \prod_j^Q \left(1 + 2\zeta_j\tau_j(j\omega) + (j\omega\tau_j)^2\right)}{\prod_m^R (1 + j\omega\tau_m) \prod_n^S (1 + 2\zeta_n\tau_n(j\omega) + (j\omega\tau_n)^2)}$$

- Cztery podstawowe elementy:
 - Stałe wzmacnienie K
 - Biegun lub zero na osi ($j\omega$) – całkowanie/różniczkowanie
 - Zero lub biegun na osi rzeczywistej ($j\omega\tau+1$) – moduł jedno-inercyjny w liczniku/mianowniku
 - Zero lub biegun sprzężony $[1+(2\zeta/\omega_n)j\omega+(j\omega/\omega_n)^2]$ – moduł oscylacyjny w liczniku/mianowniku
- Charakterystyki Bodego dla tychże elementów

Wzmocnienie

- K ma stałą charakterystykę amplitudową i fazową
 - Moduł to $20\log|K|[\text{dB}]$
 - Faza to 0 [rad]



Biegun lub zero na osi ($j\omega$) – całkowanie/różniczkowanie

- Różniczkowanie

$$G(s) = s$$

$$|G(j\omega)| = \omega$$

$$\varphi(\omega) = +\frac{\pi}{2}$$

- Całkowanie

$$G(s) = \frac{1}{s}$$

$$|G(j\omega)| = \frac{1}{\omega}$$

$$\varphi(\omega) = -\frac{\pi}{2}$$

Biegun lub zero na osi ($j\omega$) – całkowanie/różniczkowanie

- Moduł transmitancji całkowania

$$20 \log \left| \frac{1}{j\omega} \right| = -20 \log \omega [dB]$$

- Faza transmitancji całkowania

$$\varphi(\omega) = -90^\circ$$

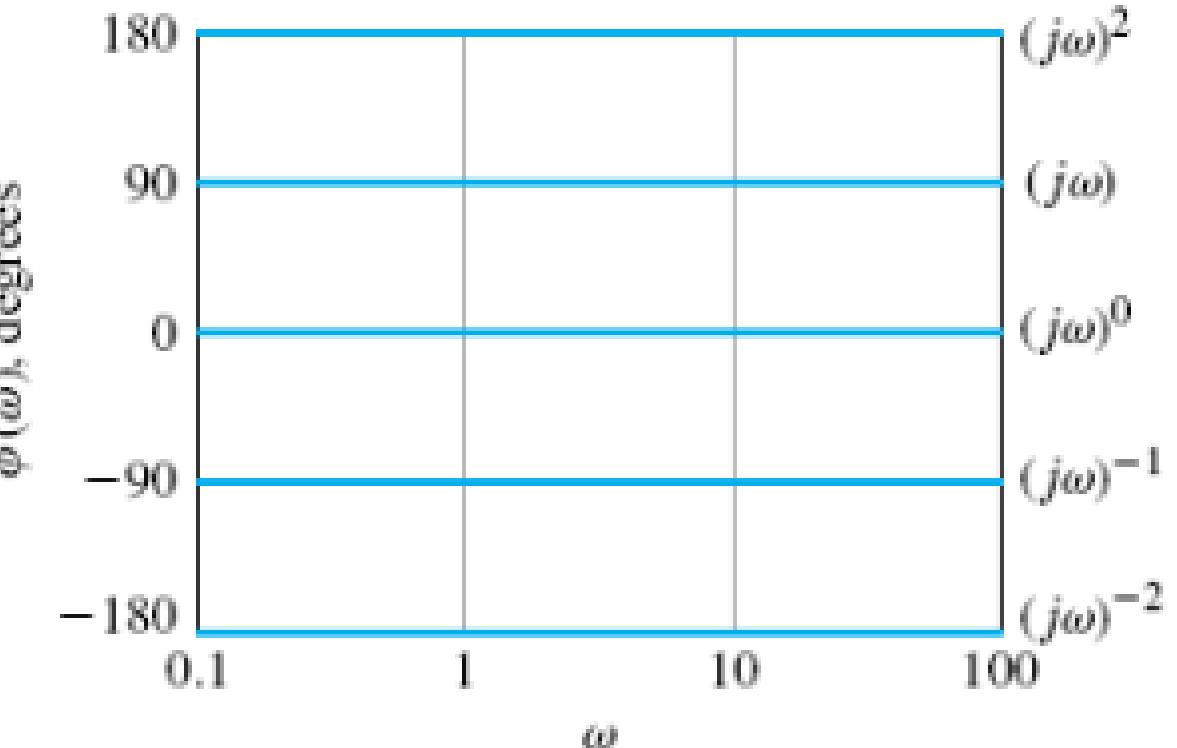
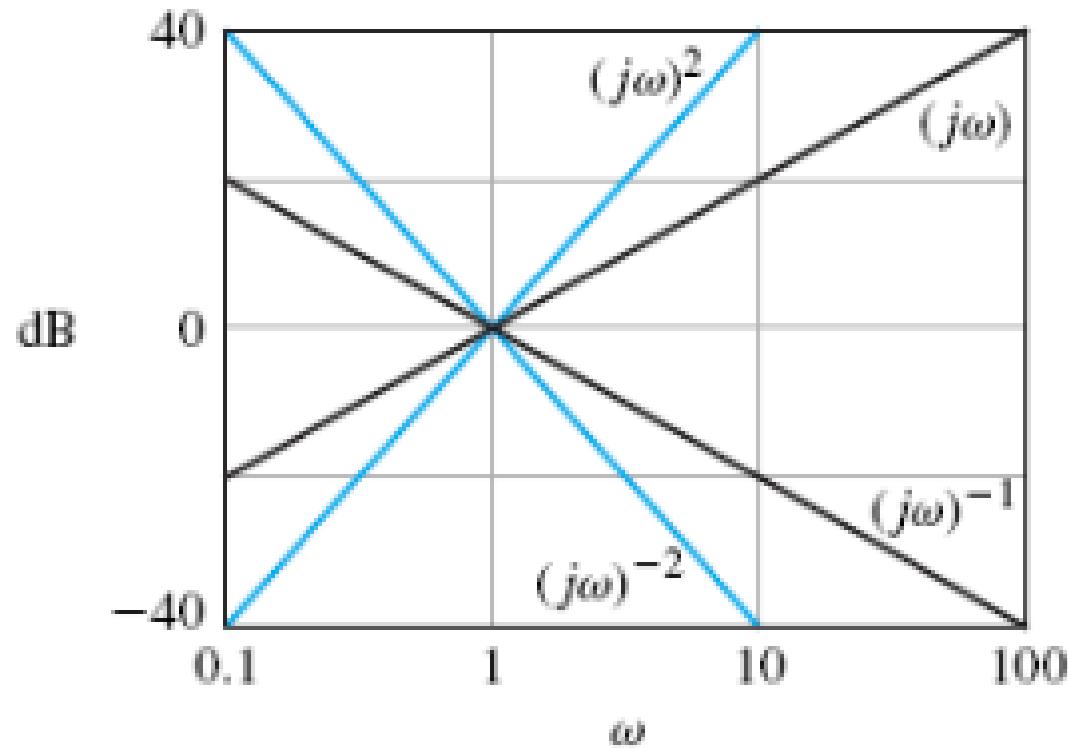
- Nachylenie charakterystyki to -20dB na dekadę dla pojedynczego całkowania
- Dla wielokrotnego całkowania mamy

$$20 \log \left| \frac{1}{(j\omega)^N} \right| = -20N \log \omega [dB]$$

- Zaś faza

$$\varphi(\omega) = -N \cdot 90^\circ$$

Biegun lub zero na osi ($j\omega$) – całkowanie/różniczkowanie



Biegun/zero rzeczywiste – układ jedno-inercyjny

- Zero

$$G(s) = \tau s + 1$$

- Moduł

$$|G(j\omega)| = \sqrt{\omega^2 \tau^2 + 1}$$

dla $\omega \ll 1/\tau$ $|G(j\omega)| \rightarrow 1$

dla $\omega \gg 1/\tau$ $|G(j\omega)| \rightarrow \omega\tau$

dla $\omega = 1/\tau$ $|G(j\omega)| \rightarrow \sqrt{2}$

- Faza

$$\varphi(\omega) = \tan^{-1} \omega\tau$$

dla $\omega \ll 1/\tau$ $\varphi(\omega) \rightarrow 0$

dla $\omega \gg 1/\tau$ $\varphi(\omega) \rightarrow \pi/2$

dla $\omega = 1/\tau$ $\varphi(\omega) = \pi/4$

- Biegun (moduł inercyjny 1-ego rzędu)

$$G(s) = \frac{1}{\tau s + 1}$$

- Moduł

$$|G(j\omega)| = \frac{1}{\sqrt{\omega^2 \tau^2 + 1}}$$

dla $\omega \ll 1/\tau$ $|G(j\omega)| \rightarrow 1$

dla $\omega \gg 1/\tau$ $|G(j\omega)| \rightarrow 1/\omega\tau$

dla $\omega = 1/\tau$ $|G(j\omega)| \rightarrow \sqrt{2}/2$

- Faza

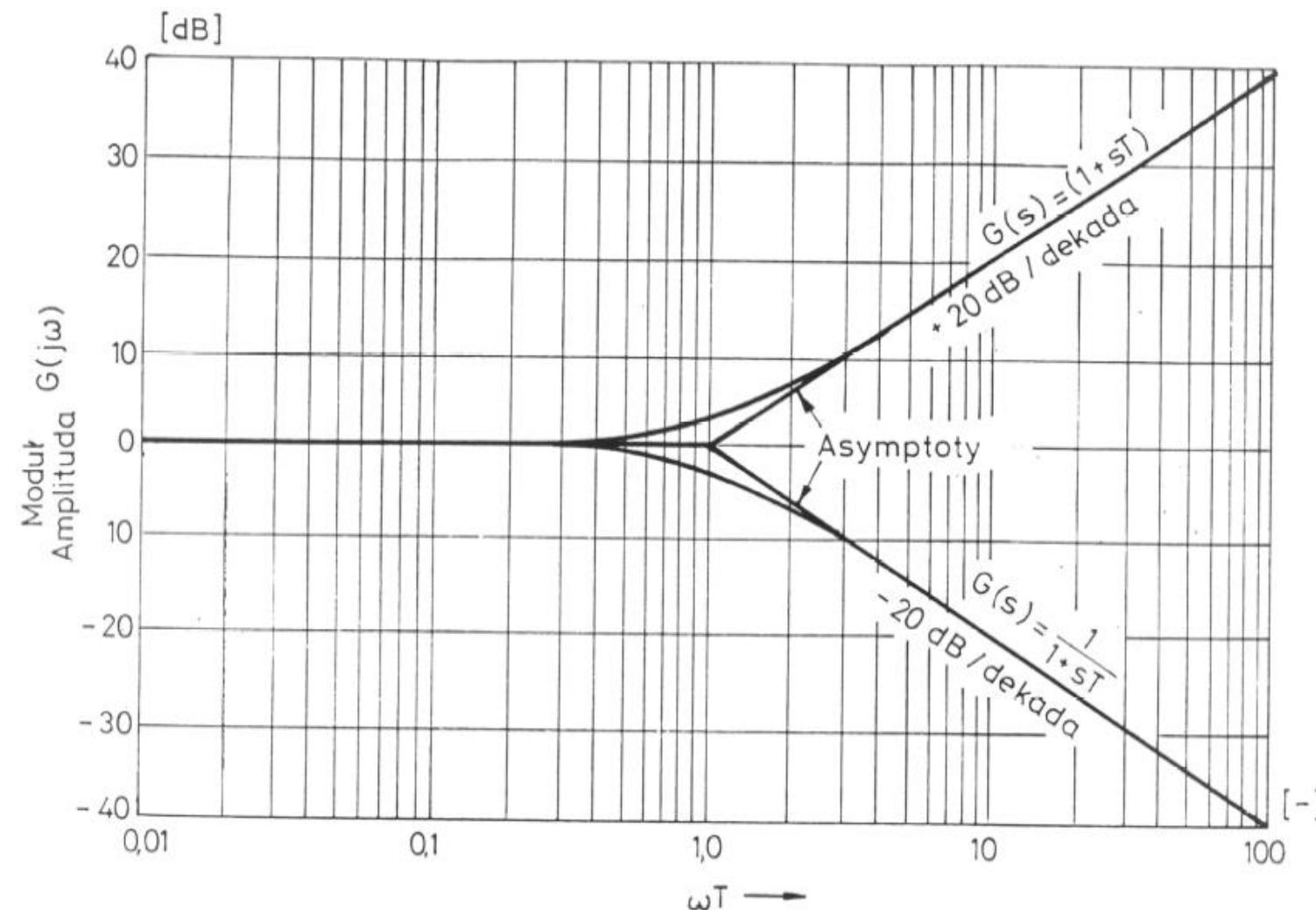
$$\varphi(\omega) = -\tan^{-1} \omega\tau = \tan^{-1} \frac{1}{\omega\tau}$$

dla $\omega \ll 1/\tau$ $\varphi(\omega) \rightarrow 0$

dla $\omega \gg 1/\tau$ $\varphi(\omega) \rightarrow -\pi/2$

dla $\omega = 1/\tau$ $\varphi(\omega) = -\pi/4$

Biegun/zero rzeczywiste – układ jedno-inercyjny



Biegun/zero rzeczywiste - właściwości

- Niskie częstotliwości $\omega \rightarrow 0$:
 - Moduł ≈ 1
 - Faza ≈ 0
- Wysokie częstotliwości $\omega \rightarrow \infty$:
 - Moduł $\approx 1/\omega\tau$
 - Faza $\approx -90^\circ$
 - Nachylenie (w dekadach) -1
- Częstotliwość odcięcia
 $\omega = \omega_b = 1/\tau$
 $AR_N(\omega = \omega_b) = -\frac{1}{\sqrt{1+1}} = 0.707$
 $\varphi(\omega = \omega_b) = \tan^{-1}(-1) = -45^\circ$
- **Krok 1.** Znajdź częstotliwość odcięcia $\omega_b = 1/\tau$. Moduł dla niej wynosi to 0.707.
- **Krok 2.** narysuj liniową stałą asymptotę ($=1$) dla niższych częstotliwości, $\omega < \omega_b$.
- **Krok 3.** Dla wyższych częstotliwości naszkicuj liniową asymptotę o nachyleniu $n-1$, z przecięciem w punkcie $\omega = \omega_b$.
- **Krok 4.** naszkicuj przebieg rzeczywisty.
- **Krok 5.** Dla $\omega = \omega_b$, $\varphi = -45^\circ$. Znajdź ten punkt i narysuj poziome asymptoty dla $\varphi = 0^\circ$ ($\omega < 0.1\omega_b$) i $\varphi = -90^\circ$ ($\omega > 10\omega_b$). Narysuj całą charakterystykę.

Zera/biegunki zespolone

- Transmitancja

$$G(s) = \frac{1}{\tau^2 s^2 + 2\xi\tau s + 1}$$

- Moduł

$$|G(j\omega)| = \frac{1}{\sqrt{(1 - \omega^2\tau^2)^2 + 4\xi^2\omega^2\tau^2}}$$

dla $\omega \ll 1/\tau$ $|G(j\omega)| \rightarrow 1$

dla $\omega \gg 1/\tau$ $|G(j\omega)| \rightarrow 1/\omega^2\tau^2$

- Faza

$$\varphi(\omega) = -\tan^{-1} \left(\frac{2\xi\omega\tau}{1 - \omega^2\tau^2} \right)$$

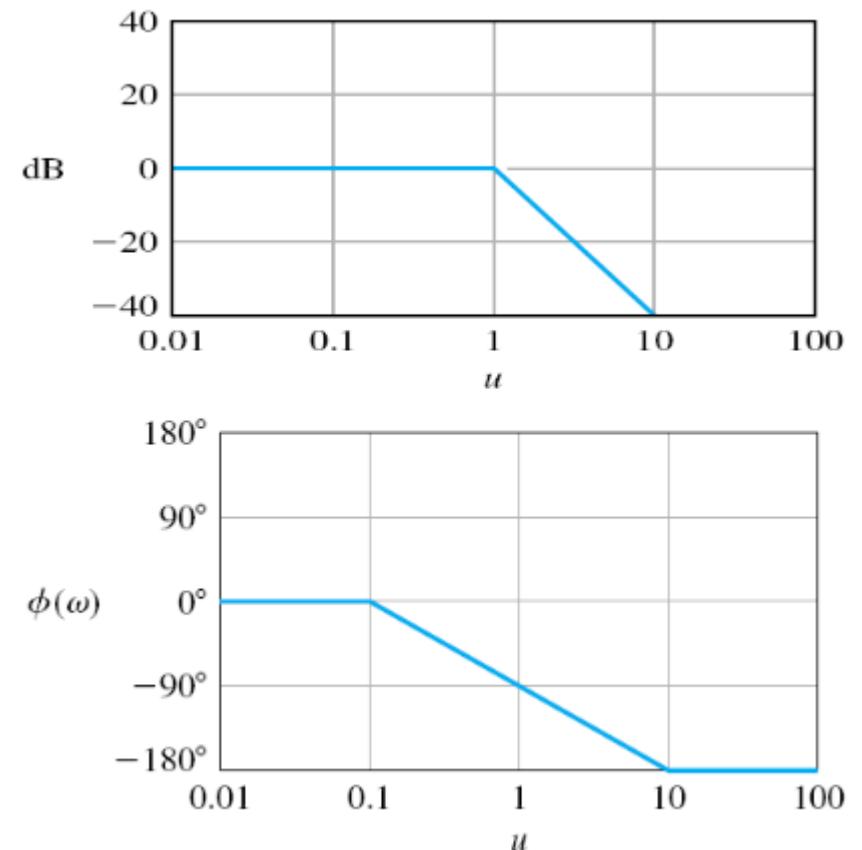
dla $\omega \ll 1/\tau$ $\varphi(\omega) \rightarrow 0$

dla $\omega \gg 1/\tau$ $\varphi(\omega) \rightarrow -\pi$

dla $\omega = 1/\tau$ $\varphi(\omega) = -\pi/2$

Zera/biegunki zespolone

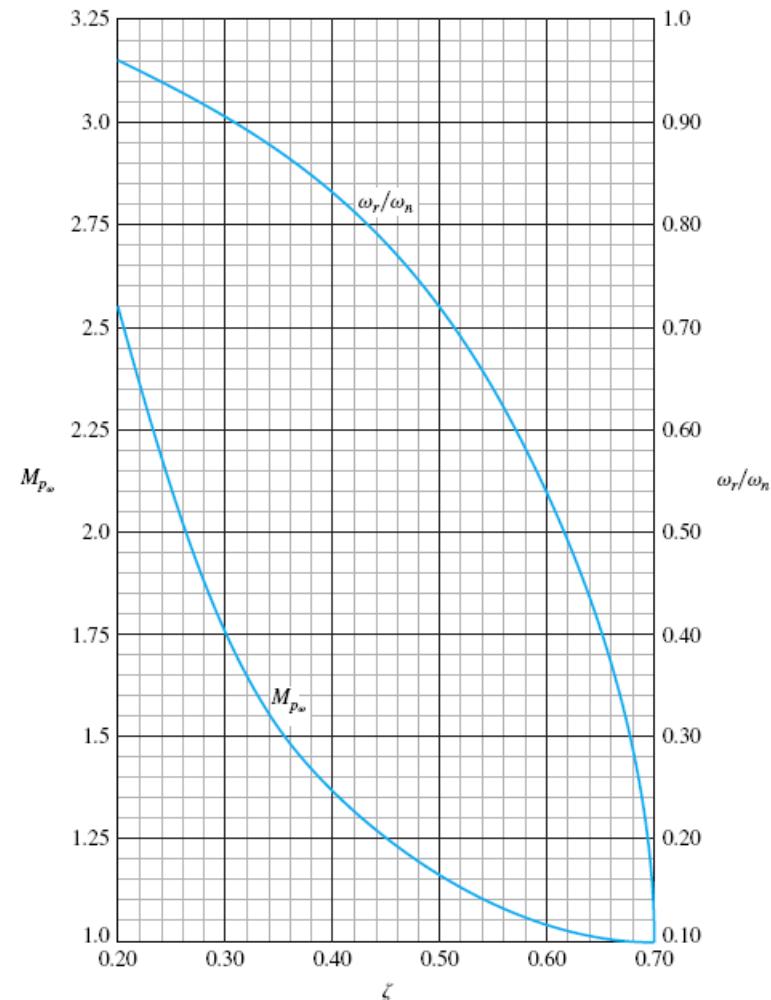
- Proste charakterystyki asymptotycznej o wartości 0 dB przecinają się dla częstotliwości odcięcia ω_b .



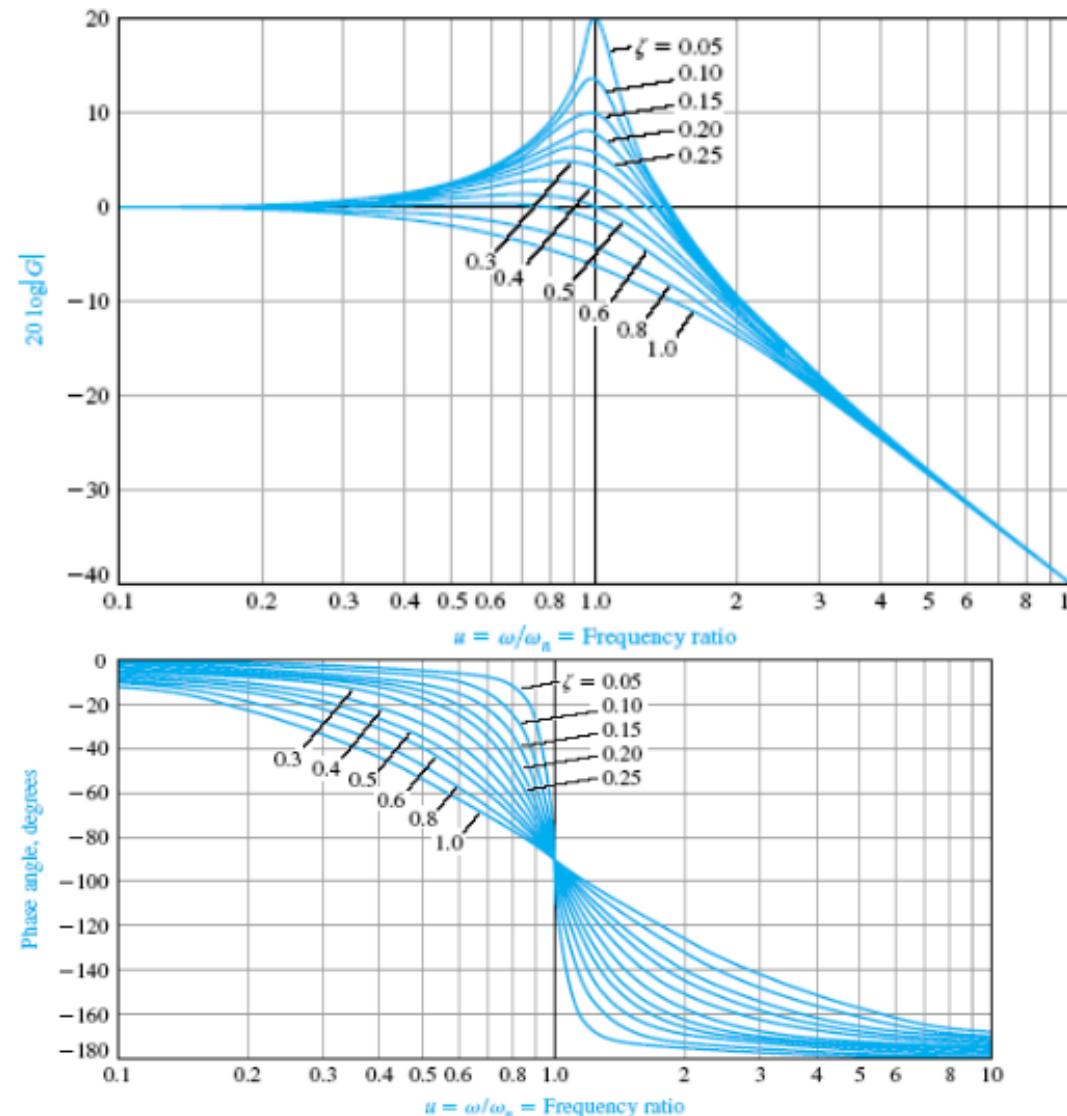
Zera/biegunki zespolone

- Częstotliwość rezonansową ω_r wyznaczamy wyznaczając punkt zerowania pochodnej modułu $[1 + j2\zeta u - u^2]^{-1}$, gdzie $u = \omega/\omega_n$
 - Częstotliwość rezonansowa wynosi
- $$\omega_r = \omega_n \sqrt{1 - 2\zeta^2}, \zeta < 0.707$$
- A maksymalna wartość modułu $|G(\omega)|$ w tym punkcie to

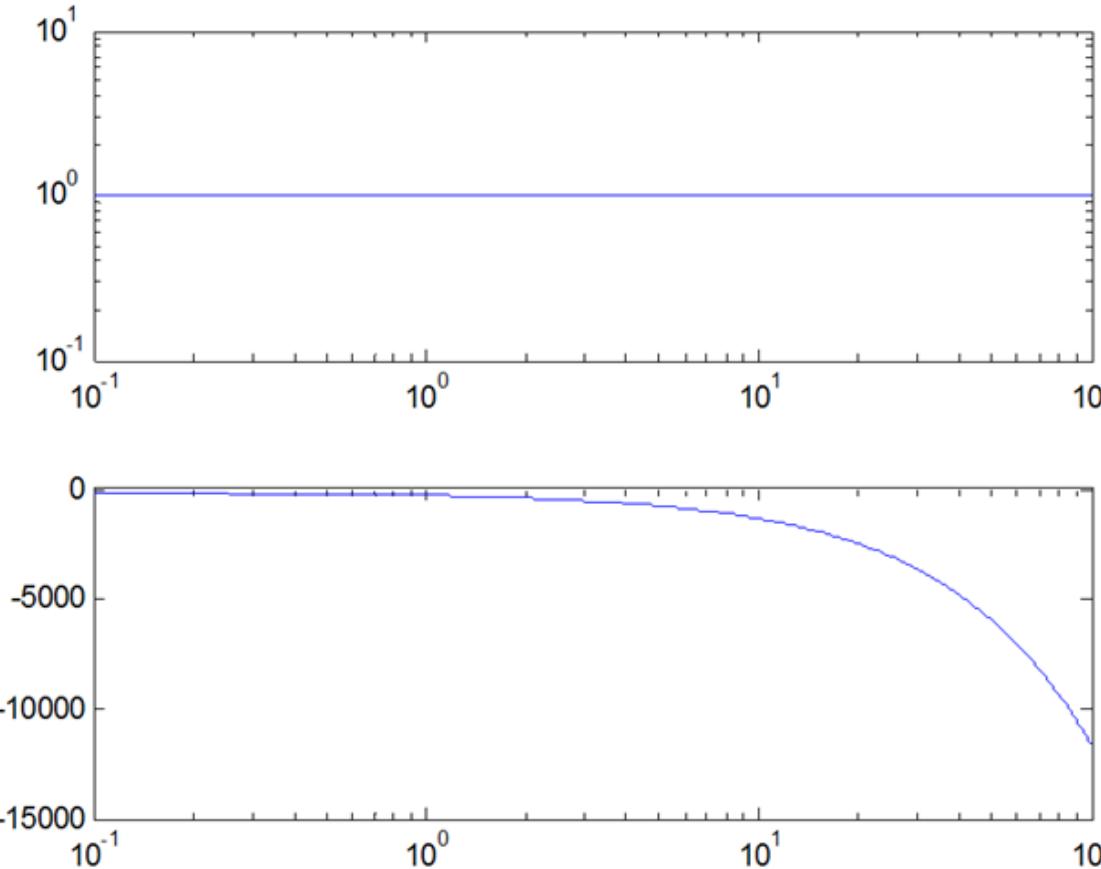
$$M_{P_\omega} = |G(\omega_r)| = [2\zeta \sqrt{1 - \zeta^2}]^{-1}, \zeta < 0.707$$



Zera/bieguny zespolone



Opóźnienie



- Charakterystyka częstotliwościowa

$$G(j\omega) = \exp(-j\omega\theta) = \cos\omega\theta - j\sin\omega\theta$$

$$\text{moduł} = |G(j\omega)| = \sqrt{\cos^2\omega\theta + \sin^2\omega\theta} = 1$$

$$\varphi = \angle G(j\omega) = \tan^{-1} \left(-\frac{\sin\omega\theta}{\cos\omega\theta} \right), \quad \varphi = -\omega\theta$$

- Nieograniczona faza!

Zera

- Zero stabilne LHP (lead)

$$G(j\omega) = j\omega T + 1$$
$$|G(j\omega)| = \sqrt{1 + \omega^2 T^2}$$

$$\omega \gg \omega_b |G(j\omega)| \rightarrow \infty$$

- Moduł dla bardzo wysokich częstotliwości ucieka do nieskończoności – fizycznie niemożliwe.
- Dodatnia faza pomiędzy 0° i 90°

$$\varphi = \tan^{-1}(\omega T)$$

Zera

- Zero niestabilne LHP (lag)

$$G(j\omega) = -j\omega T + 1$$
$$|G(j\omega)| = \sqrt{1 + \omega^2 T^2}$$

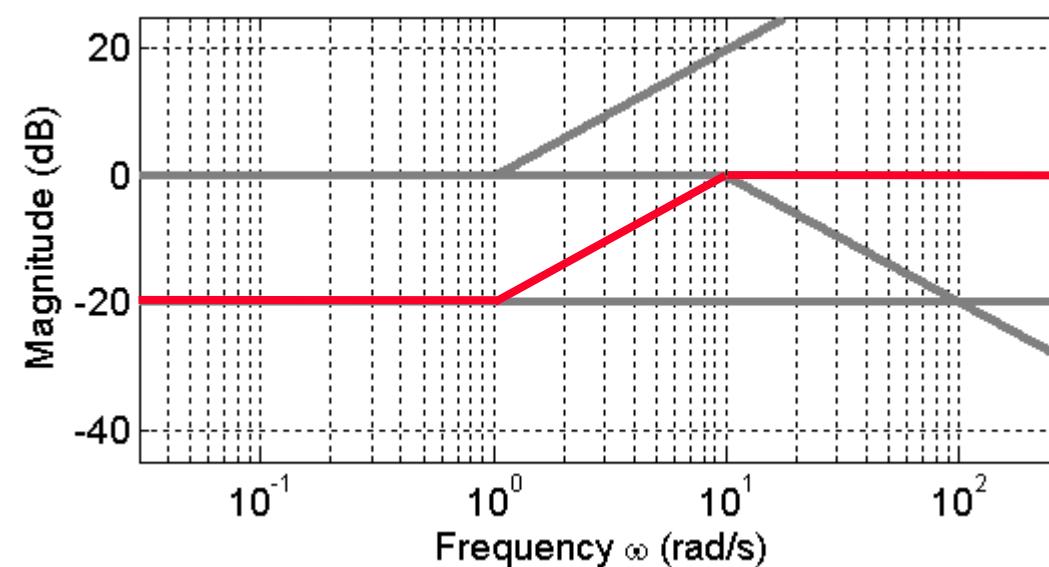
- Faza ujemna

$$\varphi = -\tan^{-1}(\omega T)$$

- Obiekt nie minimalnofazowy – tak samo jest w przypadku opóźnienia

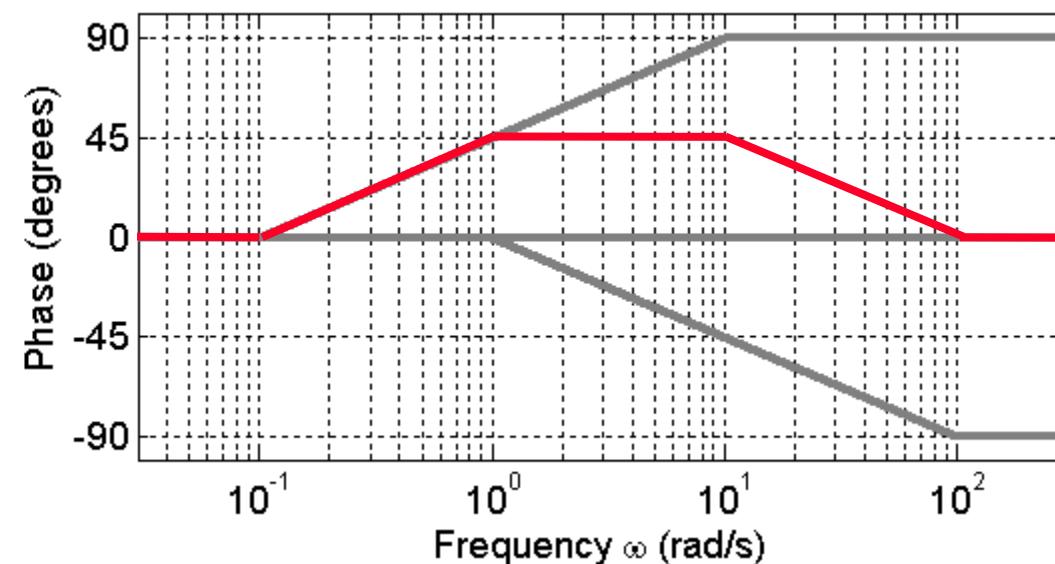
Rysowanie – przykład

$$\begin{aligned} G(s) &= \frac{s+1}{s+10} \\ &= \left[\frac{1}{10} \right] \left[\frac{1}{\frac{s}{10} + 1} \right] [s+1] \end{aligned}$$



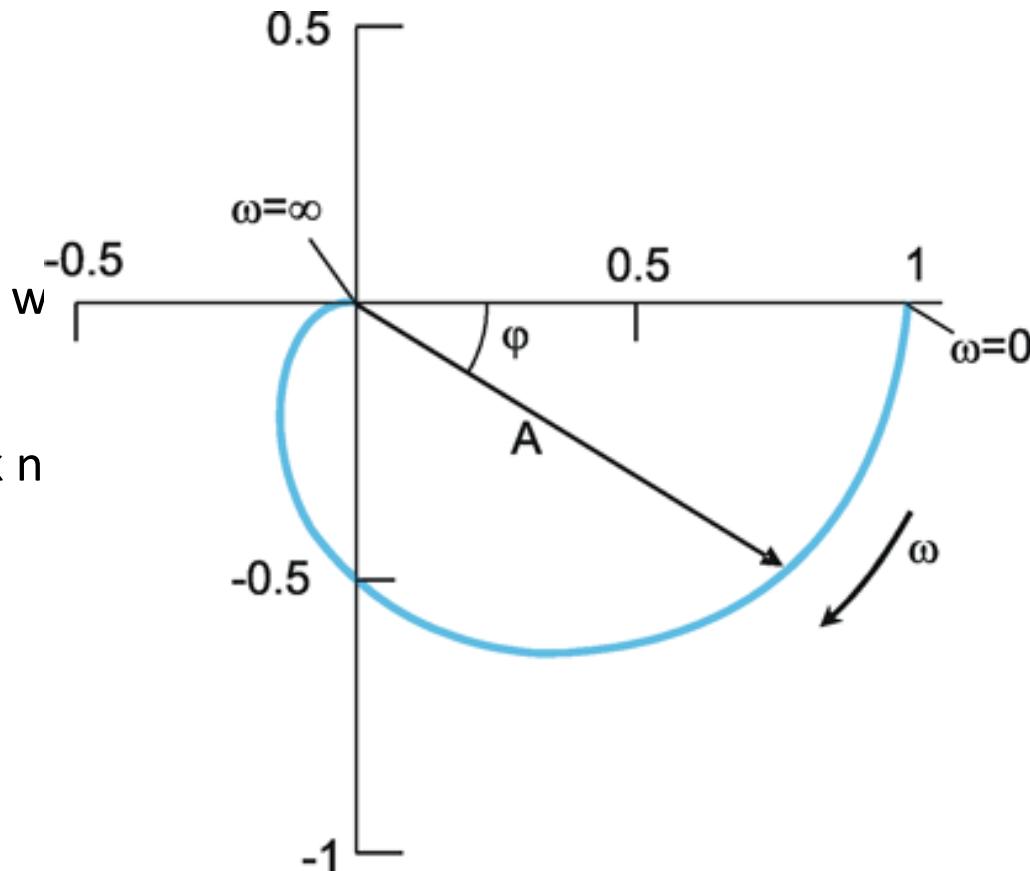
Rysowanie – przykład

$$\begin{aligned} G(s) &= \frac{s + 1}{s + 10} \\ &= \left[\frac{1}{10} \right] \left[\frac{1}{\frac{s}{10} + 1} \right] [s + 1] \end{aligned}$$



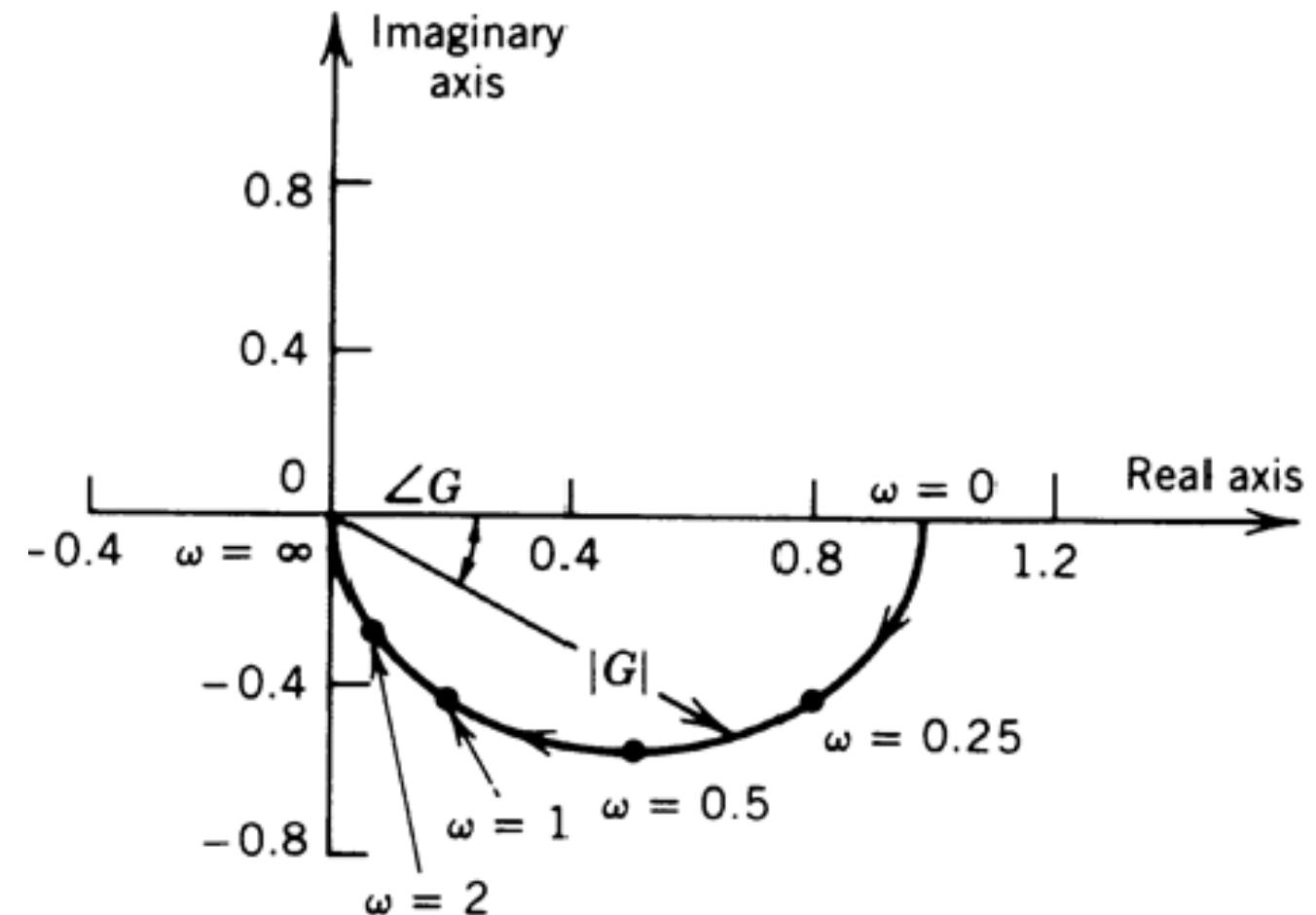
Charakterystyka Nyquista

- Wykres biegunowy $G(j\omega)$ w zależności od ω .
- Wady
 - Brak bezpośredniej informacji o częstotliwości
 - Trudniejsza w bezpośrednim rysowaniu – najpierw łatwiej w
- Zalety
 - Postać kompaktowa i wystarczająca do dalszych analiz, jak n



Charakterystyka Nyquista - przykład

- $G(s) = \frac{1}{2s+1}$
- $|G(j\omega)| = \frac{1}{\sqrt{1+4\omega^2}}$
- $\varphi = -\tan^{-1}(2\omega)$



Pobudzenie falą prostokątną (trapezową)

- Fala prostokątna może zostać opisana za pomocą szeregu Fouriera $\omega_o = 2\pi/T$

$$u(t) = \frac{4}{\pi} u_0 \left(\sin \omega_0 t + \frac{1}{3} \sin 3\omega_0 t + \frac{1}{5} \sin 5\omega_0 t + \dots \right)$$

- Zaś odpowiedź ma postać

$$\begin{aligned} y(t) = & \frac{4}{\pi} u_0 \left(|G(i\omega_0)| \sin(\omega_0 t + \varphi(\omega_0)) \right. \\ & + \frac{1}{3} |G(i3\omega_0)| \sin(3\omega_0 t + \varphi(3\omega_0)) \\ & \left. + \frac{1}{5} |G(i5\omega_0)| \sin(5\omega_0 t + \varphi(5\omega_0)) + \dots \right) \end{aligned}$$

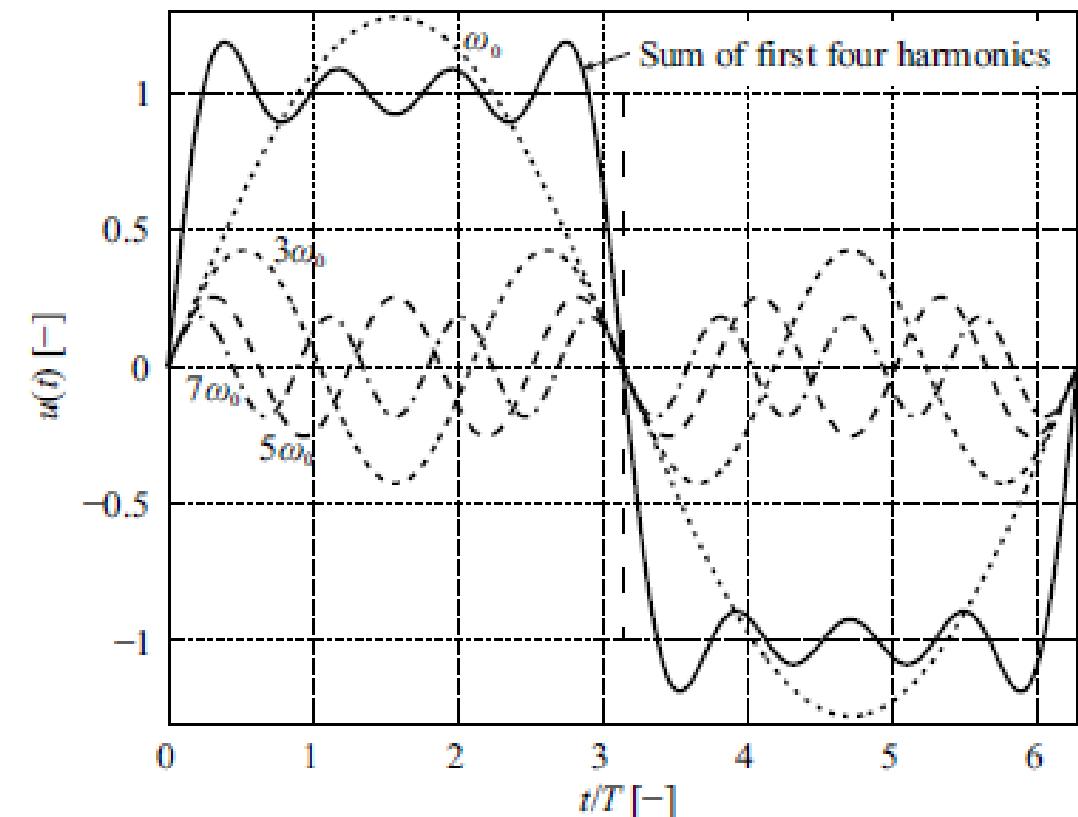
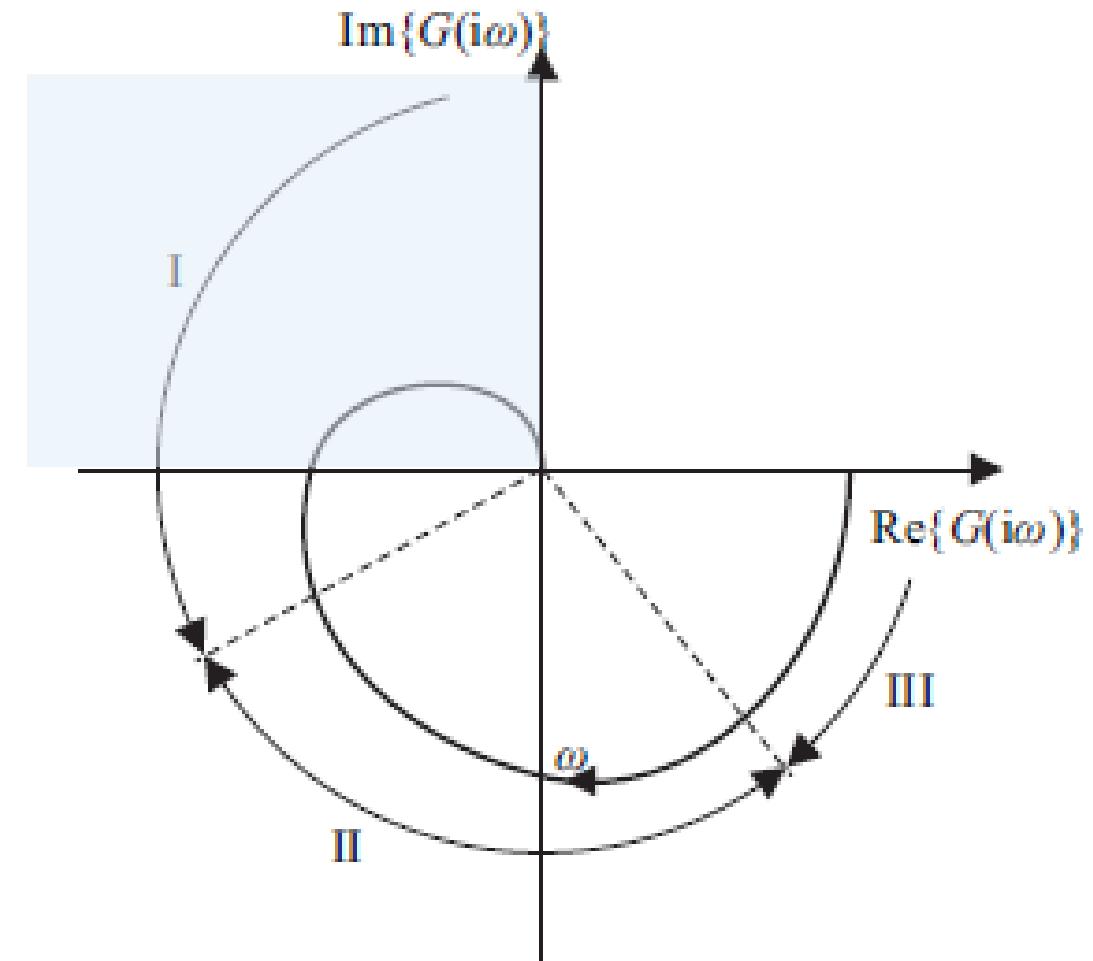


Fig. 5.3. Harmonic decomposition of a rectangular wave

Pobudzenie falą prostokątną (trapezową)

- Zaczynamy od zakresu wyższych częstotliwości
- Wyższe harmoniczne $n \geq 3$ są tłumione tak mocno, że odpowiedź odpowiada prawie że czystej sinusoidzie
- Łatwo można odczytać moduł i przesunięcie fazowe odpowiedzi
- Odczytujemy część I charakterystyki Nyquista



Pobudzenie falą prostokątną (trapezową)

- Dla średnich częstotliwości amplituda drugiej harmonicznej dla $3\omega_0$ osiąga wielkość, która nie może być pominięta

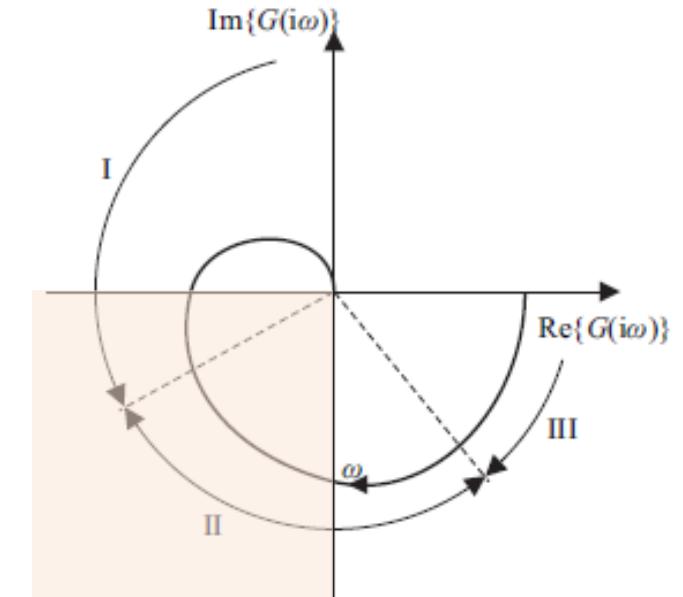
$$y(t) \approx \frac{4}{\pi} u_0 \left(|G(i\omega_0)| \sin(\omega_0 t + \varphi(\omega_0)) + \frac{1}{3} |G(i3\omega_0)| \sin(3\omega_0 t + \varphi(3\omega_0)) \right)$$

- Trzecia harmoniczna $5\omega_0$ może zostać pominięta tak jak i wyższe. Zatem można otrzymać odpowiedź dla danej częstotliwości odejmując odpowiedź dla drugiej składowej harmonicznej

$$y_{3\omega_0}(t) = \frac{4}{\pi} \frac{1}{3} u_0 |G(i3\omega_0)| \sin(3\omega_0 t + \varphi(3\omega_0))$$

od mierzonego wyjścia $y(t)$.

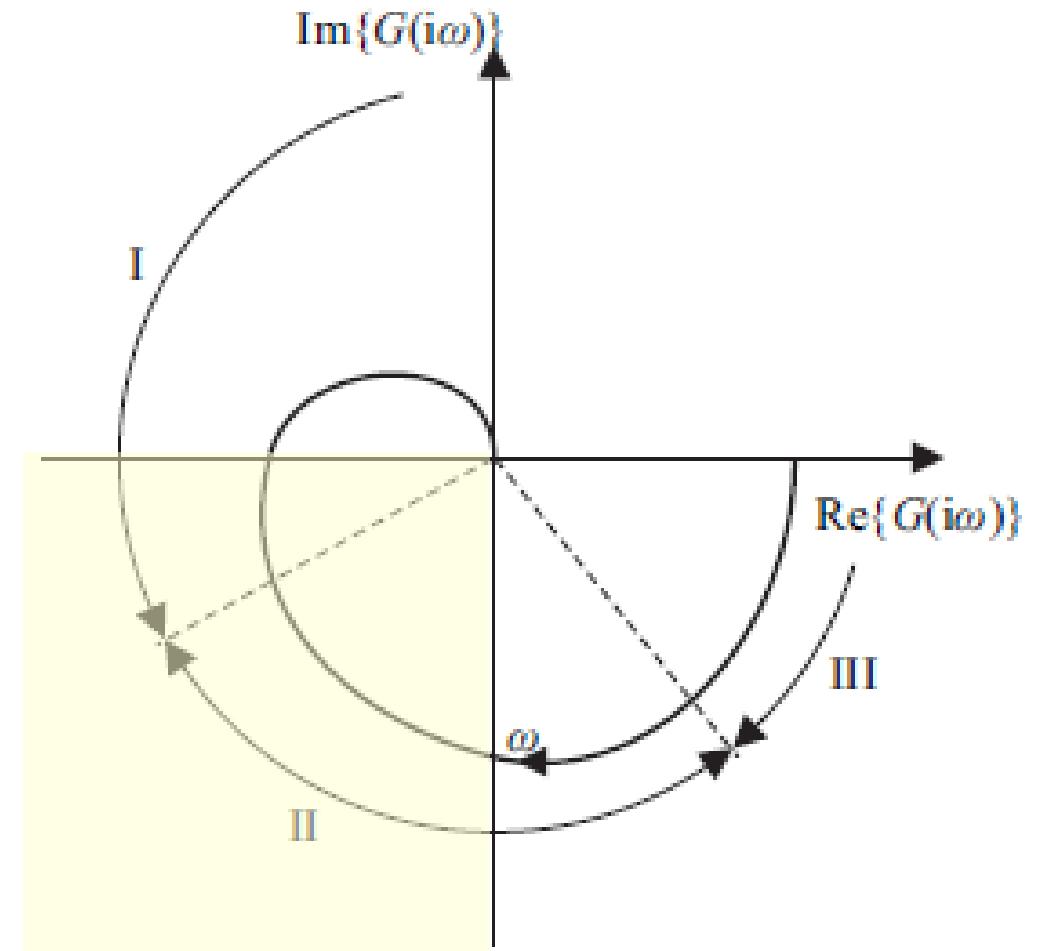
- A odpowiedź $y_{3\omega_0}$ jest znana z poprzedniego wywodu.
- Tym samym otrzymujemy strefę II.



Pobudzenie falą prostokątną (trapezową)

- Dla niskich częstotliwości odpowiedź otrzymujemy odejmując tak wiele harmonicznych jak trzeba.

$$\begin{aligned} & \frac{4}{\pi} u_0 |G(i\omega_0)| \sin(\omega_0 t + \varphi(\omega_0)) \\ &= y - u_0 \frac{1}{3} |G(i3\omega_0)| \sin(3\omega_0 t + \varphi(3\omega_0)) \\ & - u_0 \frac{1}{5} |G(i5\omega_0)| \sin(5\omega_0 t + \varphi(5\omega_0)) - \dots \end{aligned}$$



Pobudzenie falą prostokątną (trapezową)

- Metoda głównie stosowana dla wyższych zakresów częstotliwości
- Podsumowanie:
 - Łatwiej zrealizować fale prostokątną niż sinusoidalną
 - Charakterystyka urządzenia wykonawczego nie musi być liniowa
 - Dla danej amplitudy u_o fala prostokątna ma najwyższą amplitudę podstawowej harmonicznej niż dla jakichkolwiek innych pobudzeń okresowych. Zatem najlepszy stosunek do zakłóceń.
 - Ze względu na charakter przekształcenia Fouriera czas przełączenia od $+u_o$ do $-u_o$ powinien być nie większy niż:

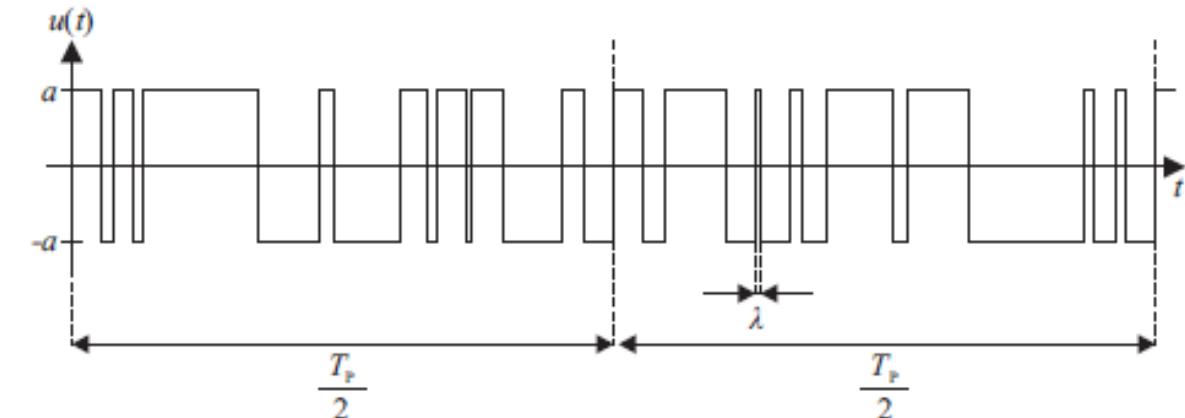
$$T_1^* < \frac{1.1}{\omega_{\max}} \text{ resp. } T_1^* < \frac{0.5}{\omega_{\max}}$$

- Gdzie ω_{\max} to najwyższa częstotliwość w interesującym zakresie

Pobudzenie z wieloma harmonicznymi

- Binarny sygnał o wielu częstotliwościach
- Cel: maksymalizacja amplitud dla poszczególnych składowych częstotliwościowych przy minimalizacji amplitudy właściwego pobudzenia

$$\left. \begin{aligned} |G(i\omega_v)| &= \frac{1}{u_{0v}} \sqrt{a_{yyv}^2 + b_{yyv}^2} \\ \varphi(\omega_v) &= \arctan \frac{a_{yyv}}{b_{yyv}} \end{aligned} \right\}$$



Sześć (6) harmonicznych

$$u_o = 0.585a$$

$$N = 256$$

$$\omega_o, 2\omega_o, 4\omega_o, 8\omega_o, 16\omega_o, 32\omega_o$$

Sygnały z ciągłą zmianą częstotliwości

Equation

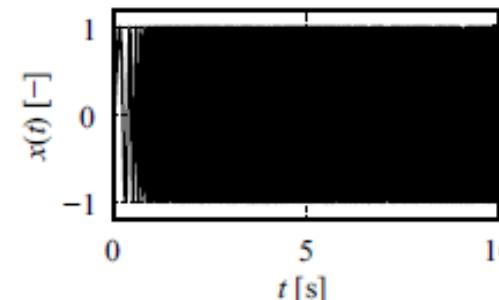
$$x(t) = \sin(2\pi f(t) t)$$

with

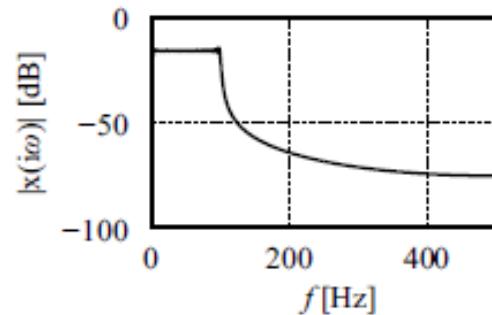
$$f(t) = f_0 + \frac{f_1 - f_0}{2T} t$$

Linear Sweep

Time Response



Fourier Transform



Equation

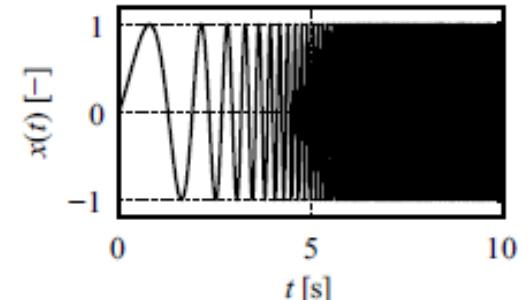
$$x(t) = \sin(2\pi f(t) t)$$

with

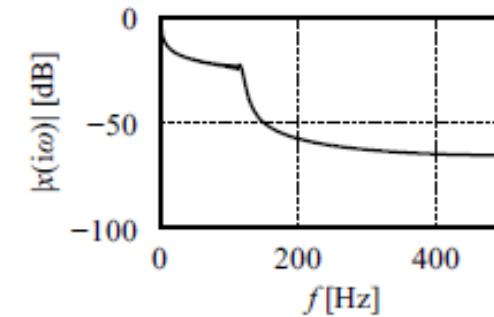
$$f(t) = f_0 \left(\frac{f_1}{f_0}\right)^t$$

Exponential Sweep

Time Response



Fourier Transform



Stosowane do analizy w teorii obwodów i identyfikacji sieci (ogólnie telekomunikacja)



MI

Metody Identyfikacji

wykład #5-b

1. *Metoda korelacyjna*

Metoda koreacyjna dla charakterystyk częstotliwościowych

- Funkcja autokorelacji

$$R_{uu}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} u(t)u(t + \tau)dt$$

- Funkcja korelacji wzajemnej

$$\begin{aligned} R_{uy}(\tau) &= E\{u(t)y(t + \tau)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} u(t)y(t + \tau)dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} u(t - \tau)y(t)dt . \end{aligned}$$

- Obie funkcje są ze sobą powiązane

$$R_{uy}(\tau) = \int_0^{\infty} g(t')R_{uu}(\tau - t')dt'$$

- Ważne nie tylko dla sygnałów stochastycznych ale i okresowych

Metoda koreacyjna dla charakterystyk częstotliwościowych

- Możliwość wyznaczenia odpowiedzi impulsowej (poprzez znalezienie funkcji autokorelacji oraz korelacji wzajemnej)
- Zaś na podstawie odpowiedzi impulsowej (transformata Fouriera) można znaleźć charakterystyki częstotliwościowe
- Ze względu na poniższe cechy **można bezpośrednio wyznaczyć moduł i fazę!**
- Dla pobudzenia sinusoidalnego

$$u(t) = u_o \sin \omega_o t, \text{ o częstotliwości } \omega_o = 2\pi/T_p$$

- Funkcja autokorelacji przyjmuje postać

$$R_{uu}(\tau) = \frac{2u_0^2}{T_0} \int_0^{\frac{T_0}{2}} \sin(\omega_0 t + \alpha) \sin(\omega_0(t + \tau) + \alpha) dt = \frac{u_0^2}{2} \cos \omega_0 \tau$$

Metoda korelacyjna dla charakterystyk częstotliwościowych

- Funkcja korelacji wzajemnej sygnału testowego i odpowiedzi na sygnał testowy

$$y(t) = u_0 |G(i\omega_0)| \sin(\omega_0 t - \varphi(\omega_0))$$

- prowadzi do postaci

$$\begin{aligned} R_{uy}(\tau) &= |G(i\omega_0)| \frac{2u_0^2}{T_p} \int_0^{\frac{T_p}{2}} \sin \omega_0(t-\tau) \sin(\omega_0 t - \varphi(\omega_0)) dt \\ &= |G(i\omega_0)| \frac{u_0^2}{2} \cos(\omega_0 \tau - \varphi(\omega_0)). \end{aligned}$$

- Ze względu na okresowość całkowanie można ograniczyć do połowy okresu.
- Ostatecznie otrzymujemy:

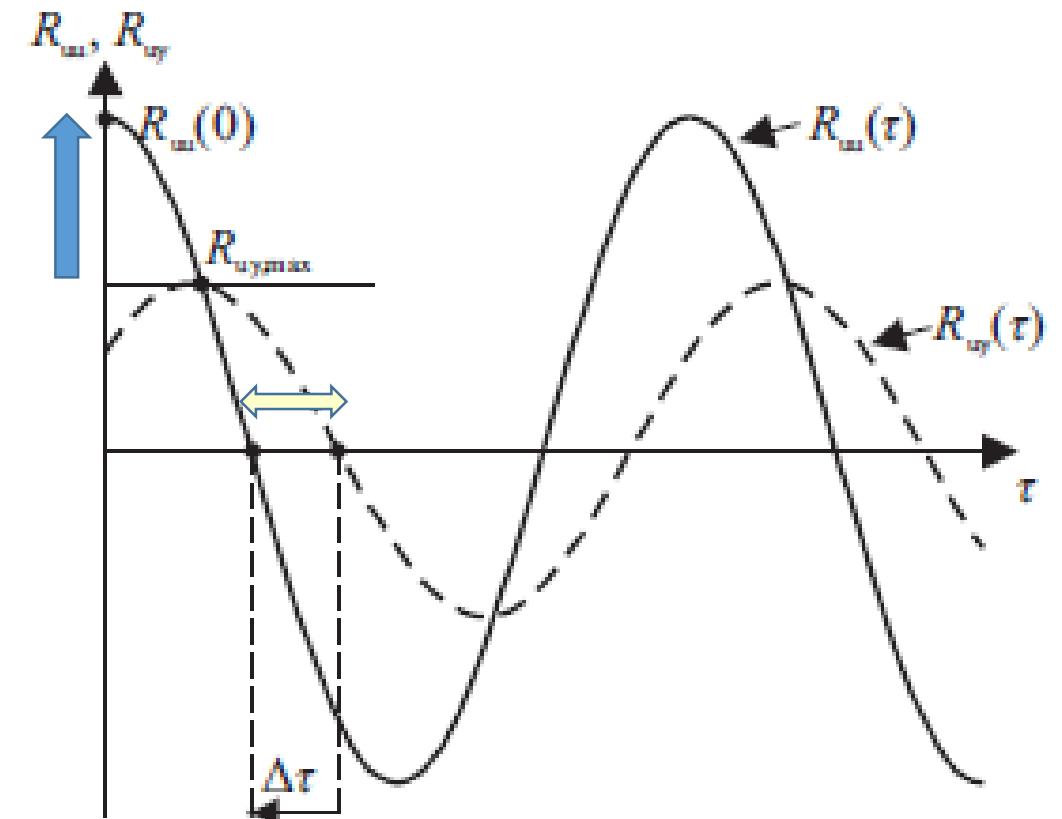
$$R_{uy}(\tau) = |G(i\omega_0)| R_{uu}\left(\tau - \frac{\varphi(\omega_0)}{\omega_0}\right)$$

Metoda koreacyjna dla charakterystyk częstotliwościowych

- Rysując obie korelacje w funkcji czasu otrzymujemy:
 - Moduł charakterystyki to stosunek korelacji wzajemnej w punkcie τ do funkcji autokorelacji w momencie $\tau - \varphi(\omega_o)/\omega_o$

$$|G(i\omega_0)| = \frac{R_{uy}(\tau)}{R_{uu}\left(\tau - \frac{\varphi(\omega_0)}{\omega_0}\right)} = \frac{R_{uy,\max}}{R_{uu}(0)} = \frac{R_{uy}\left(\frac{\varphi(\omega_0)}{\omega_0}\right)}{R_{uu}(0)}$$

- A faza jest określana z opóźnienia $\Delta\tau$
 $\varphi(\omega_o) = -\omega_o\Delta\tau$



Metoda koreacyjna dla charakterystyk częstotliwościowych

- Metoda nie jest ograniczona dla sygnałów sinusoidalnych
- Mogą być sygnały o wielu harmonicznych dopóki korelacje są sinusoidalne
- W obecności zakłóceń należy rozpatrywać dłuższe przebiegi – wiele okresów
 - Błąd zanika jeśli zakłócenie stochastyczne nie jest skorelowane z wejściem
 - Dotyczy to również przebiegów okresowych, tyle że muszą mieć częstotliwość różną od ω_o

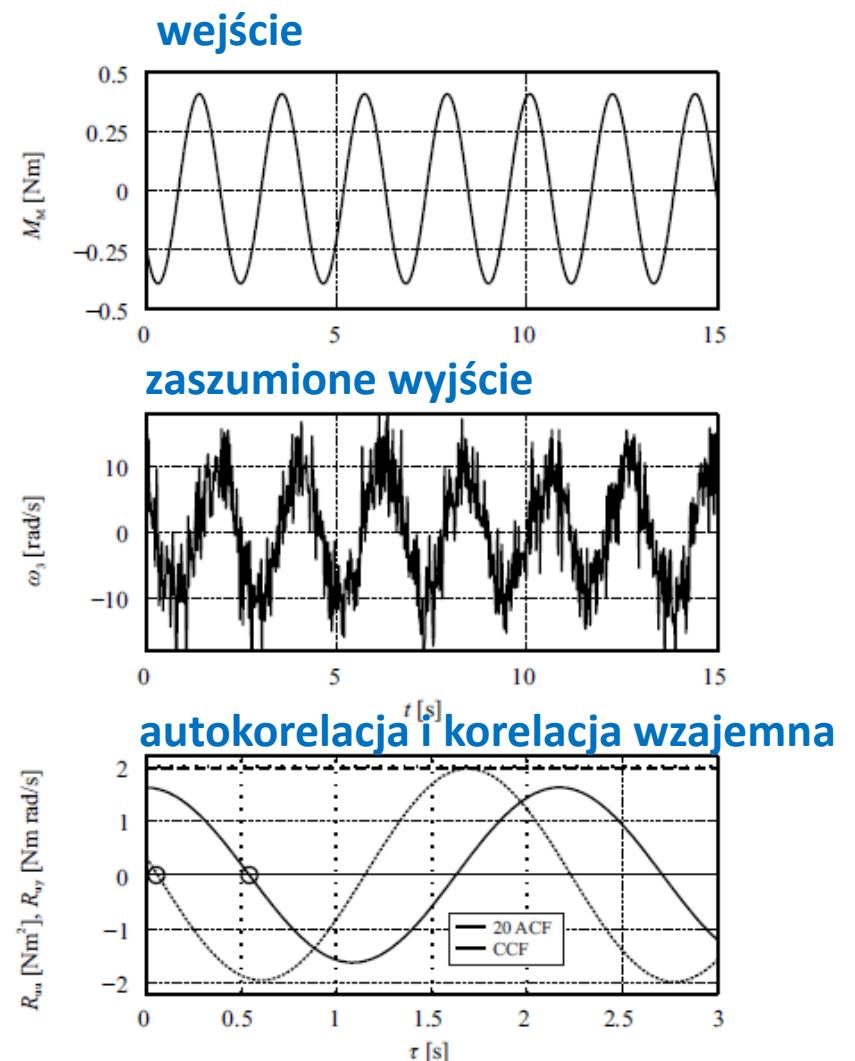
Przykład oscylatora trzech mas

- Otrzymujemy moduł

$$|G(i\omega_0)|_{\omega=2.8947 \text{ rad/s}} = \frac{R_{uy,\max}}{R_{uu}(0)} = \frac{1.99 \text{ Nm} \frac{\text{rad}}{\text{s}}}{0.081 \text{ Nm}^2} = 24.49 \frac{\text{rad}}{\text{s Nm}}$$

- oraz fazę

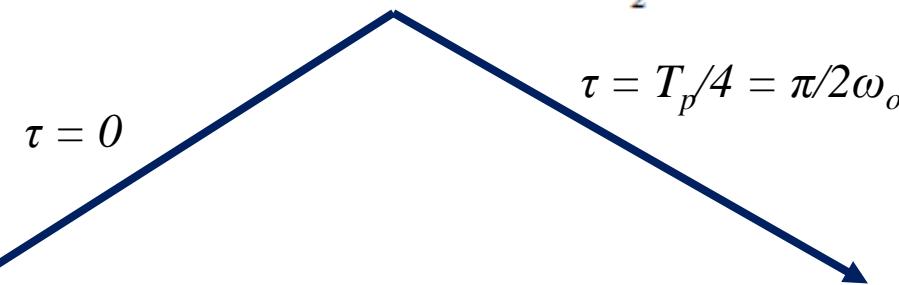
$$\begin{aligned} \varphi(i\omega)|_{\omega=2.8947 \text{ rad/s}} &= -\Delta \tau \omega = (0.542 \text{ s} - 0.053 \text{ s}) 2.8947 \frac{\text{rad}}{\text{s}} \\ &= -1.41 \text{ rad} = -81.1^\circ. \end{aligned}$$



Korelacja ortogonalna

- Można wyznaczyć część rzeczywistą i urojoną charakterystyki z funkcji korelacji wzajemnej

$$|G(i\omega_0)| \cos(\omega_0 \tau - \varphi(\omega_0)) = \frac{R_{uy}(\tau)}{\frac{u_0^2}{2}}$$

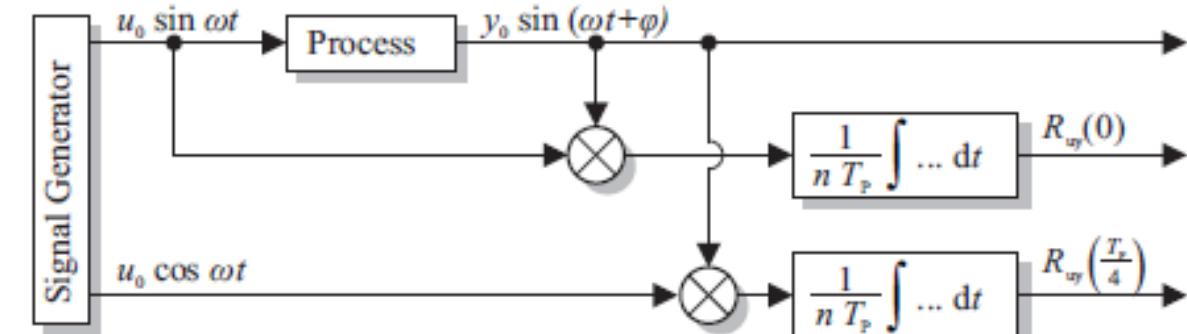


$$\operatorname{Re}\{G(i\omega_0)\} = |G(i\omega_0)| \cos(\varphi(\omega_0)) = \frac{R_{uy}(0)}{\frac{u_0^2}{2}}$$

$$R_{uy}(0) = \frac{u_0^2}{2} \operatorname{Re}\{G(i\omega_0)\} = \frac{u_0}{n T_p} \int_0^{n T_p} y(t) \sin \omega_0 t \, dt$$

$$\operatorname{Im}\{G(i\omega_0)\} = |G(i\omega_0)| \sin(\varphi(\omega_0)) = \frac{R_{uy}\left(\frac{\pi}{2\omega_0}\right)}{\frac{u_0^2}{2}}$$

$$R_{uy}\left(\frac{T_p}{4}\right) = \frac{u_0^2}{2} \operatorname{Im}\{G(i\omega_0)\} = -\frac{u_0}{n T_p} \int_0^{n T_p} y(t) \cos \omega_0 t \, dt$$



- Zatem wystarcza wyznaczyć funkcję korelacji wzajemnej jedynie w dwu punktach

Korelacja ortogonalna

- Dzięki ortogonalności funkcji trygonometrycznych otrzymujemy podstawowe zależności:

$$\begin{aligned} R_{uy}(0) &= \frac{1}{nT_p} u_0 y_0 \left(\int_0^{nT_p} \sin^2 \omega_0 t \cos \varphi dt + \underbrace{\int_0^{nT_p} \sin \omega_0 t \cos \omega_0 t \sin \varphi dt}_{=0} \right) \\ &= \frac{y_0}{u_0} \frac{u_0^2}{2} \cos \varphi = |G(i\omega_0)| \cos \varphi \frac{u_0^2}{2} = \text{Re}\{G(i\omega_0)\} \frac{u_0^2}{2}. \end{aligned}$$

$$\begin{aligned} R_{uy}\left(\frac{T_p}{4}\right) &= \frac{1}{nT_p} \int_0^{nT_p} u_0 \cos \omega_0 t y_0 \sin(\omega_0 t + \varphi) dt \\ &= \text{Im}\{G(i\omega_0)\} \frac{u_0^2}{2}. \end{aligned}$$

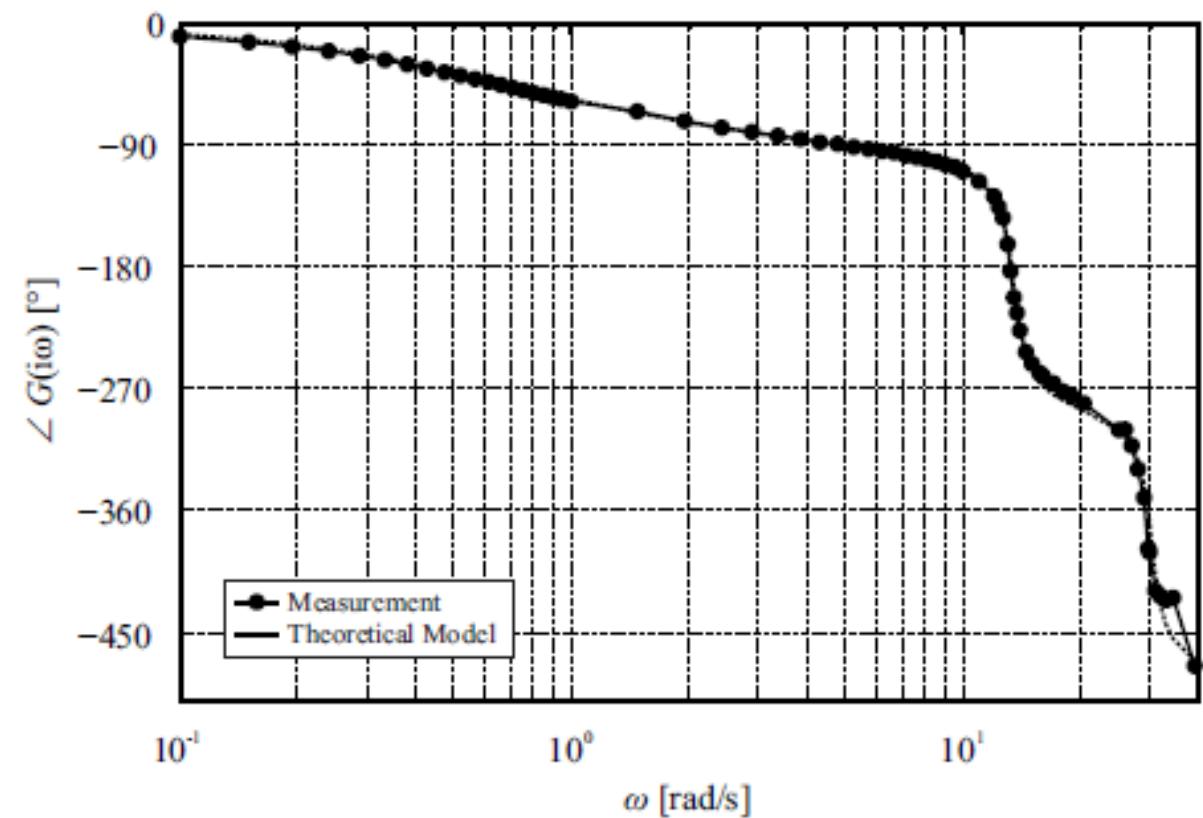
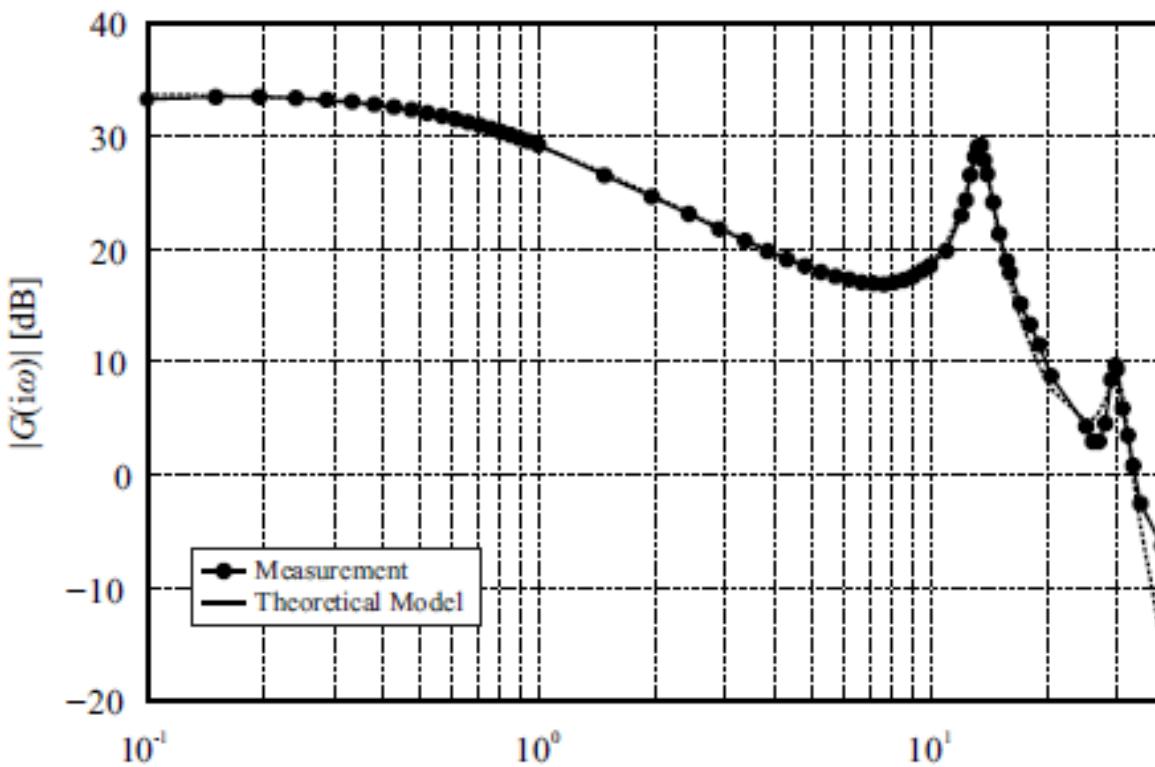
- Amplituda i faza otrzymana jest w postaci

$$|G(i\omega_0)| = \sqrt{\text{Re}^2\{G(i\omega_0)\} + \text{Im}^2\{G(i\omega_0)\}}$$

$$\varphi(\omega_0) = \arctan \frac{\text{Im}\{G(i\omega_0)\}}{\text{Re}\{G(i\omega_0)\}}.$$

- Metoda klasyczna – nie jest zakłócana częstotliwościami innymi od ω_o

Przykład: oscylator trzech mas

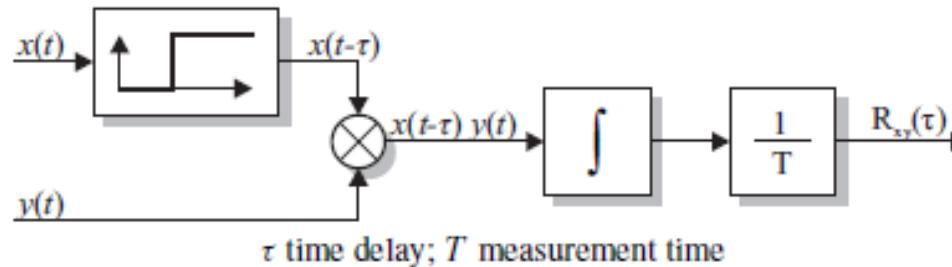


Metoda korelacyjna

- Praktyczne zastosowanie w celu ograniczenia czasu doświadczenia:
 - Niskie i średnie częstotliwości → analiza Fourierowska
 - Wysokie częstotliwości → metoda korelacyjna
- Co dalej?

Pobudzanie sygnałami zawierającymi wiele harmonicznych!!!

Estymacja funkcji korelacji wzajemnej



$$\begin{aligned}
 R_{xy}(\tau) &= E\{x(t)y(t + \tau)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)y(t + \tau)dt \\
 &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t - \tau)y(t)dt
 \end{aligned}$$

$\longrightarrow R_{xy}(\tau) = -R_{yx}(\tau)$

$$\begin{aligned}
 R_{yx}(\tau) &= E\{y(t)x(t + \tau)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} y(t)x(t + \tau)dt \\
 &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} y(t - \tau)x(t)dt
 \end{aligned}$$

Wpływ długości okresu T eksperymentu jest istotny!!!

Estymacja funkcji korelacji wzajemnej

- Wpływy zakłóceń zanikają wraz z przyrostem długości okresu T
- Wariancja estymaty korelacji wzajemnej zanika odwrotnie proporcjonalnie do czasu T
- Estymata korelacji wzajemnej jest nieobciążona, jeśli zakłócenie addytywne na wejściu i wyjściu jest odpowiednio niezależne od wejścia i wyjścia
 - Wariancja estymaty jest wtedy równa sumie wariancji obu zakłóceń
 - **Oba zakłócenia muszą być od siebie niezależne**

Estymacja funkcji autokorelacji

- Estymata jest nieobciążona
- Jeśli sygnał jest zakłócony

$$x(t) = x_o(t) + n(t)$$

- otrzymujemy

$$R_{xx}(\tau) = R_{xoxo}(\tau) + R_{nn}(\tau)$$

Estymaty funkcji korelacji wzajemnej i autokorelacji

- Obie funkcje korelacji i korelacji wzajemnej dla obiektów liniowych o odpowiedzi impulsowej $g(t)$ są estymowane w sensie minimalnokwadratowym przy następujących warunkach:
 - Sygnały wejściowy i wyjściowy jest **stacjonarny**
 - Wartość oczekiwana wejścia **wynosi zero**
 - Zakłócenie jest **stacjonarne i nieskorelowane** z wejściem
 - Analogicznie jak niezależnie zakłócone addytywnie jest wejście i wyjście, ale oba zakłócenia muszą być **nieskorelowane wzajemnie**



MI

Metody Identyfikacji

wykład #6

1. *Sygnały pseudolosowe*
2. *Identyfikacja modeli parametrycznych*

Biały szum jako wejście

- Funkcja autokorelacji

$$R_{uu}(\tau) = S_{u0} \delta(\tau)$$

- Funkcja korelacji wzajemnej

$$R_{uy}(\tau) = S_{u0} \int_0^{\infty} g(t')\delta(\tau - t')dt' = S_{u0} g(\tau)$$

- Odpowiedź impulsowa

$$g(\tau) = \frac{1}{S_{u0}} R_{uy}(\tau)$$

Niestety biały szum jest nierealizowalny!!!

Szum szerokopasmowy

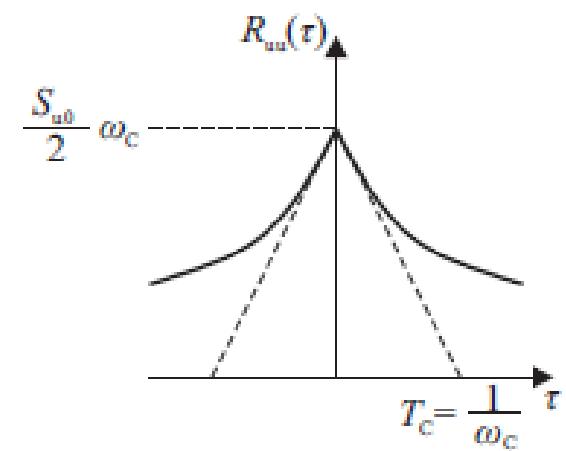
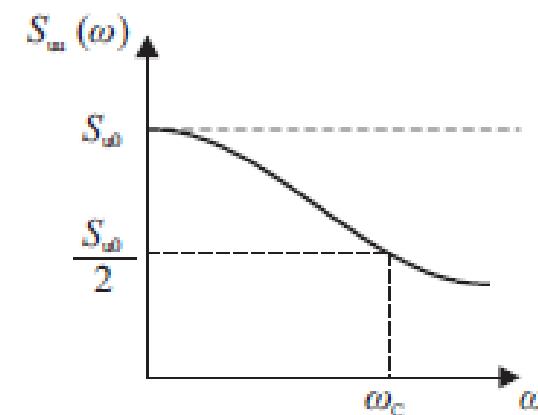
- Filtrowanie *białego* szumu

$$S_{uu}(\omega) = |G_F(i\omega)|^2 S_{u0}$$

- Dla filtru jedno-inercyjnego

$$\begin{aligned} R_{uu}(\tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} |G_F(i\omega)|^2 S_{u0} e^{i\omega t} d\omega \\ &= \frac{1}{\pi} \int_0^{\infty} \frac{S_{u0}}{1 + T_c^2 \omega^2} \cos \omega \tau d\omega \\ &= \frac{1}{2} S_{u0} \omega_c e^{-\omega_c |\tau|}. \end{aligned}$$

- Błąd wynikający ze skończonego pasma przenoszenia maleje wraz ze wzrostem pasma (częstotliwości granicznej)
- Niestety jednocześnie rośnie błąd wynikający z zakłócenia
- Tym samym **pasmo nie może być za duże**



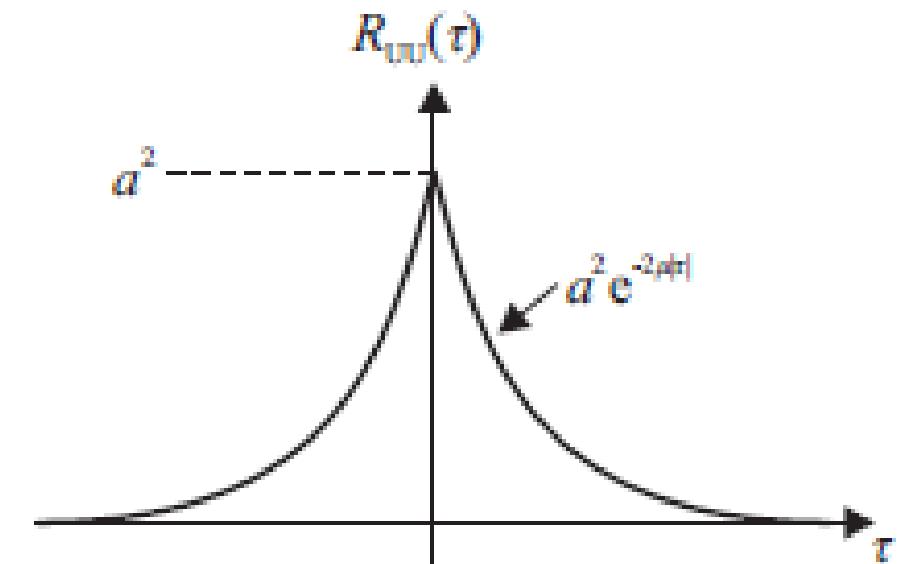
Wykorzystanie naturalnych szumów jako wejście

- Muszą być spełnione następujące warunki:
 - Stacjonarny
 - Pasmo przenoszenia musi być wyższe od największej interesującej częstotliwości
 - Gęstość widmowa musi być większa niż zakłócenia na wyjściu – aby skrajnie nie wydłużać eksperymentu
 - Nie może być skorelowany z innymi zakłóceniami
 - Otwarta pętla
 - Nie w trybie ręcznym

Binarne sygnały losowe

- Ciągły sygnał binarny – RBS
Continuous-time Random Binary Signal
- Dwa stany: +a i -a
- Zmiana występuje w dowolnie wybranym czasie
- Zalety:
 - Prosta generacja sygnału
 - Proste wyznaczenie funkcji korelacji wzajemnej
 - **Wysoki stosunek gęstości widmowej w zależności od amplitudy sygnału**

$$\begin{aligned} E\{u(t)u(t+\tau)\} &= a^2(P(0) + P(2) + \dots) - a^2(P(1) + P(3) + \dots) \\ &= a^2 e^{-\mu|\tau|} \left(1 - \frac{\mu\tau}{1!} + \frac{(\mu\tau)^2}{2!} \pm \dots\right) \\ &= a^2 e^{-2\mu|\tau|}. \end{aligned}$$

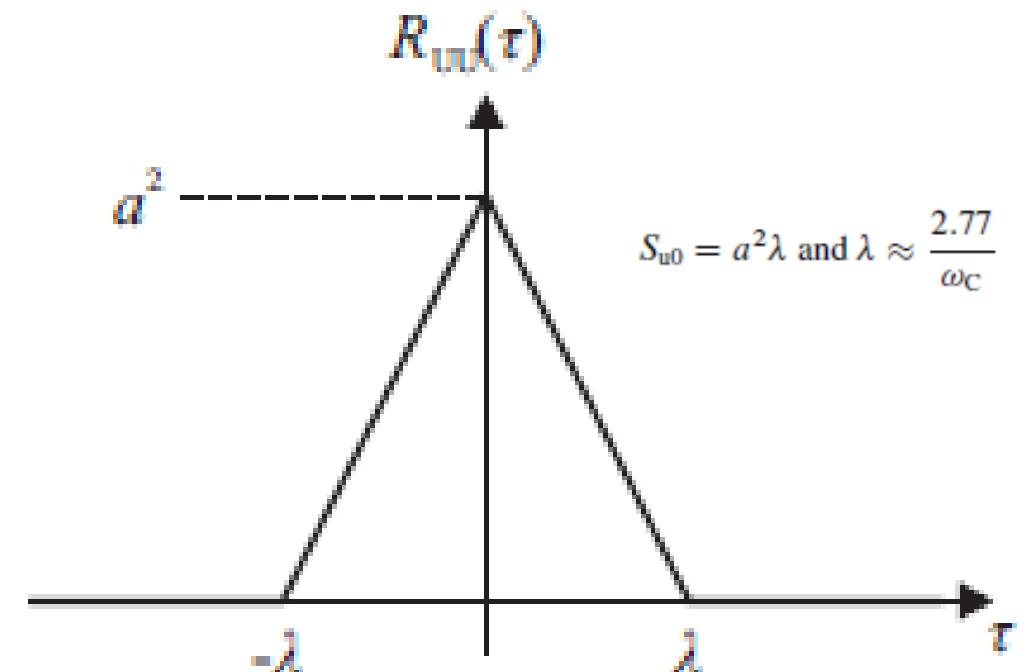


Binarne sygnały losowe

- Dyskretny sygnał binarny – DRBS
Discrete Random Binary Signal
- Dwa stany: +a i -a
- Zmiana występuje z określonym okresem próbkowania
- Zalety:
 - Nad wyraz prosta generacja sygnału
 - Proste wyznaczenie funkcji korelacji wzajemnej
 - Wysoki stosunek gęstości widmowej w zależności od amplitudy sygnału

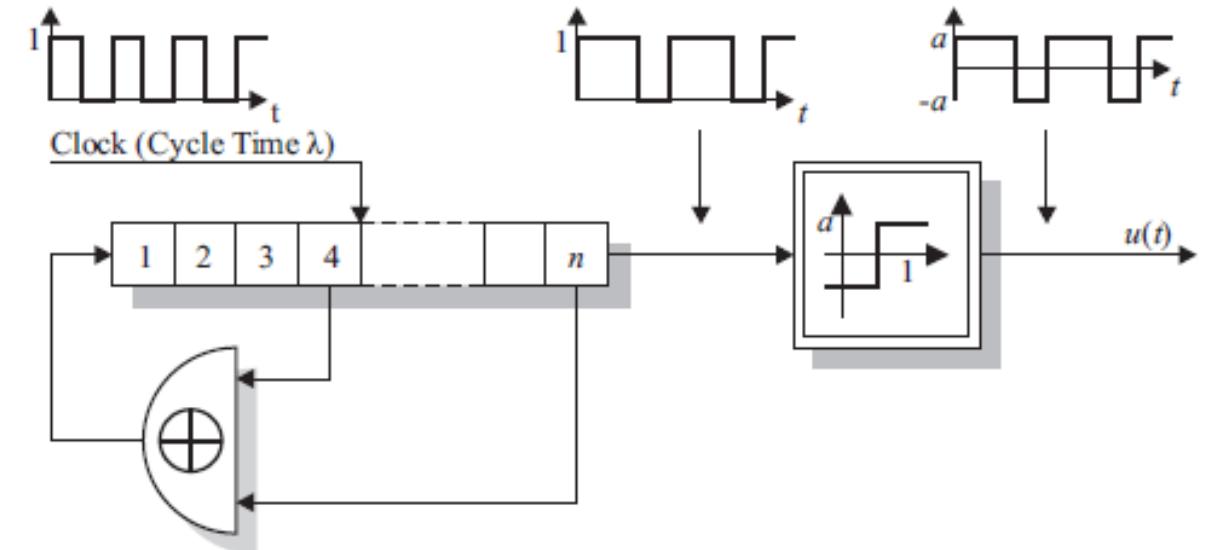
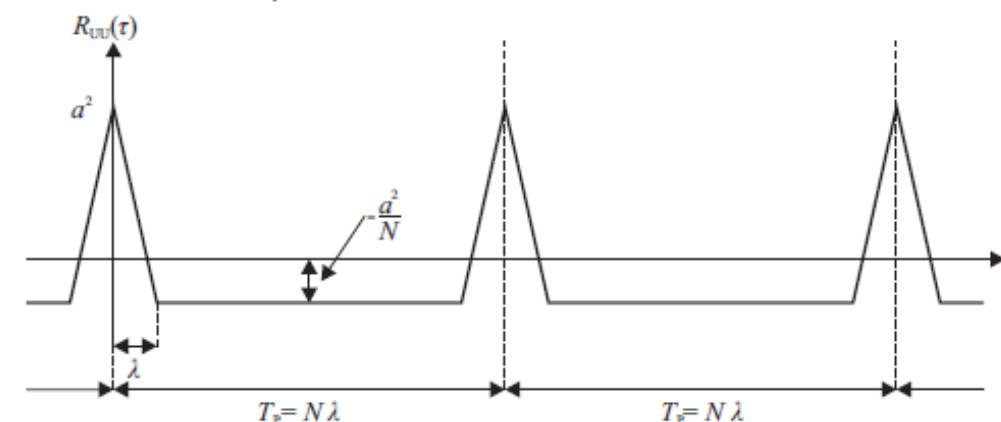
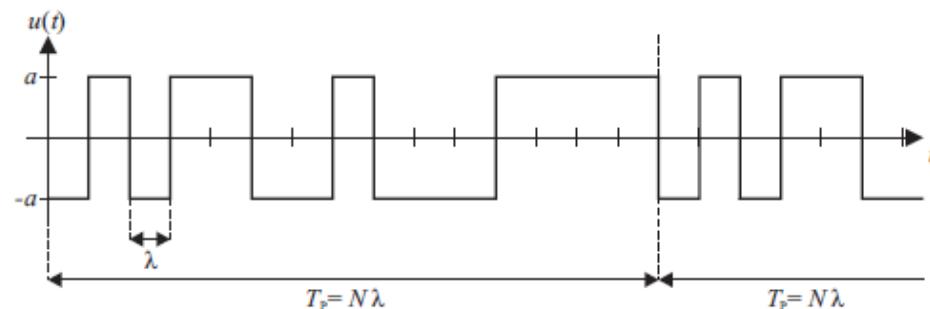
$$R_{uu}(\tau) = \begin{cases} a^2 \left(1 - \frac{|\tau|}{\lambda}\right) & \text{for } |\tau| < \lambda \\ 0 & \text{for } |\tau| \geq \lambda \end{cases}$$

λ - okres próbkowania



PRBS – pseudolosowy sygnał binarny

- PRBS – Pseudo Random Binary Signal
- Długość: $N\lambda = (2^n - 1)\lambda$
- Średnia = a/N



Okresowa funkcja autokorelacji

PRBS – przykładowe realizacje

No. of Stages	Feedback Law	Length
2	1 XOR 2	3
3	1 XOR 3 or 2 XOR 3	7
4	1 XOR 4 or 3 XOR 4	15
5	2 XOR 5 or 3 XOR 5	31
6	1 XOR 6 or 5 XOR 6	63
7	1 XOR 7 or 3 XOR 7 or 4 XOR 7 or 6 XOR 7	127
8	1 XOR 2 XOR 7 XOR 8	255
9	4 XOR 9 or 5 XOR 9	511
10	3 XOR 10 or 7 XOR 10	1023
11	2 XOR 11 or 9 XOR 11	2047

PRBS - dobór

- Zasady wyboru parametrów:

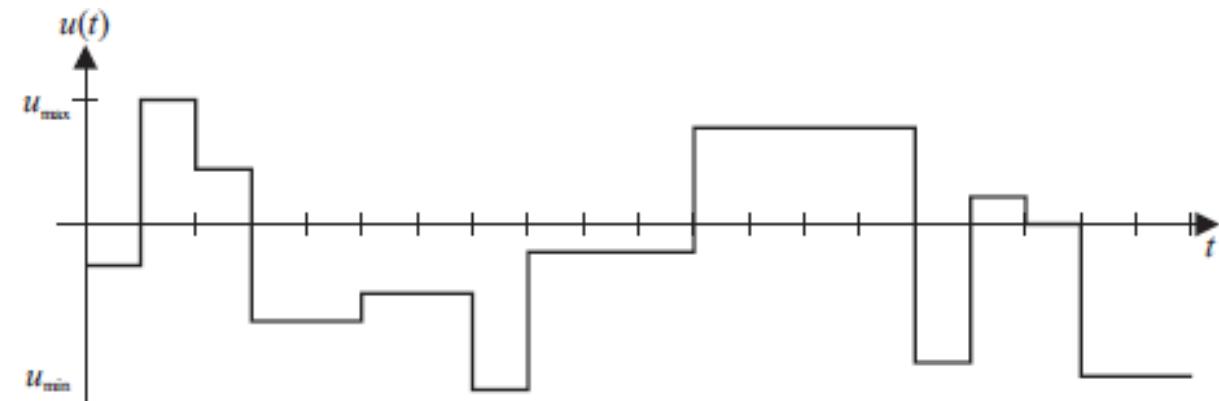
- Amplituda największa z możliwych, tak aby ograniczyć wpływ zakłóceń, ale trzeba uwzględnić ograniczenia procesowe/technologiczne na sterowanie i wyjście
- Im większy okres próbkowania tym lepiej, aby częstotliwość odcięcia $\omega_c = 1/\lambda$ nie była za mała

Sugerowany wybór:

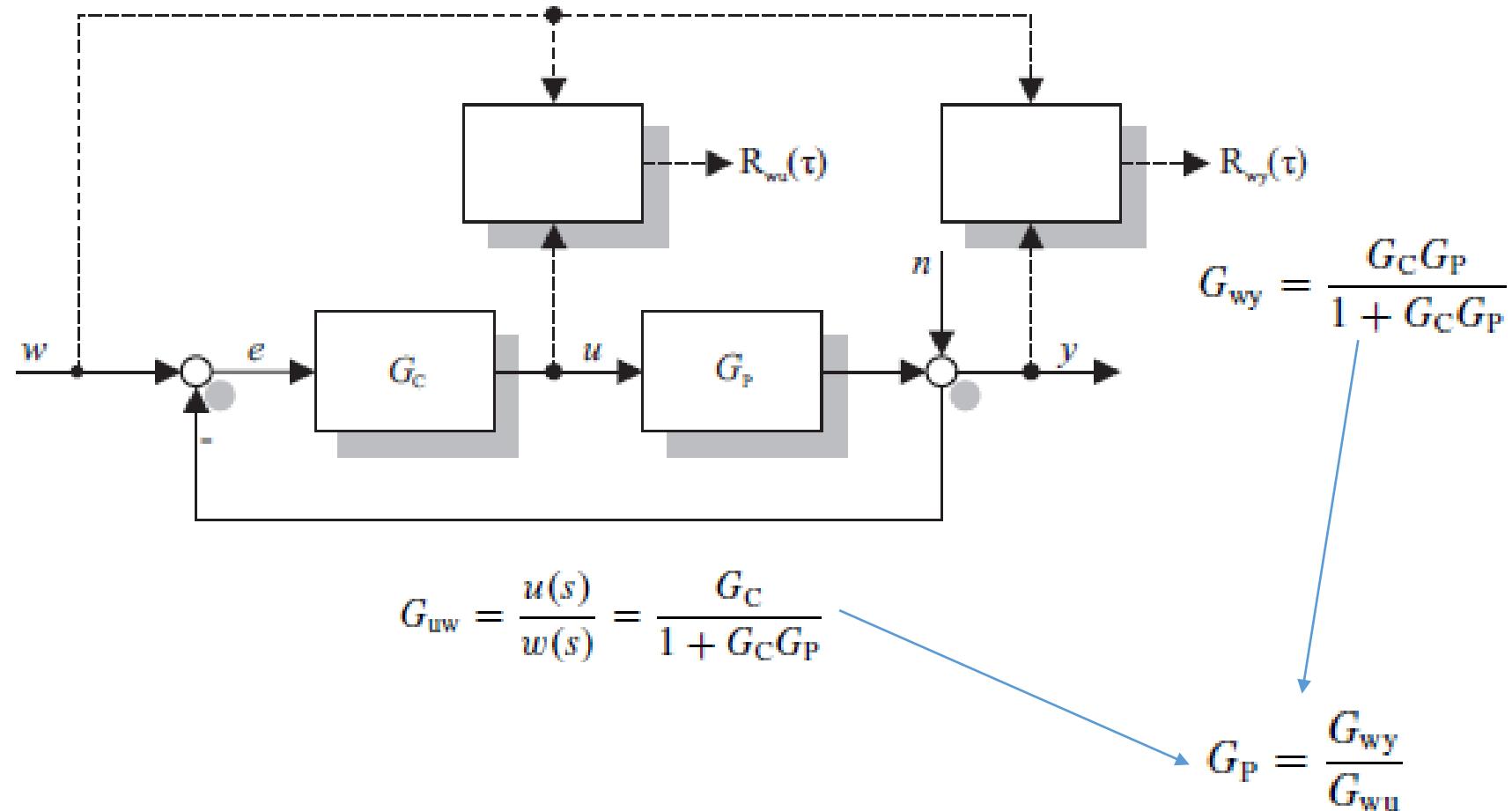
- $\lambda \leq T_i/5$ (T_i – najmniejsza stała czasowa procesu)
- $\lambda \leq T_{\text{dom}}/10$ (T_{dom} – dominująca stała czasowa procesu)
- Długość okresu PRBS nie powinna być dłuższa od czasu ustalenia procesu T_{95}
 - $T_p \approx 1.5 T_{95}$

PRBS zmodulowany amplitudowo

- Dla obiektów nieliniowych
- Wystarczająco pobudzony sygnał wejściowy



Analiza korelacyjna w pętli zamkniętej



MNK - historia

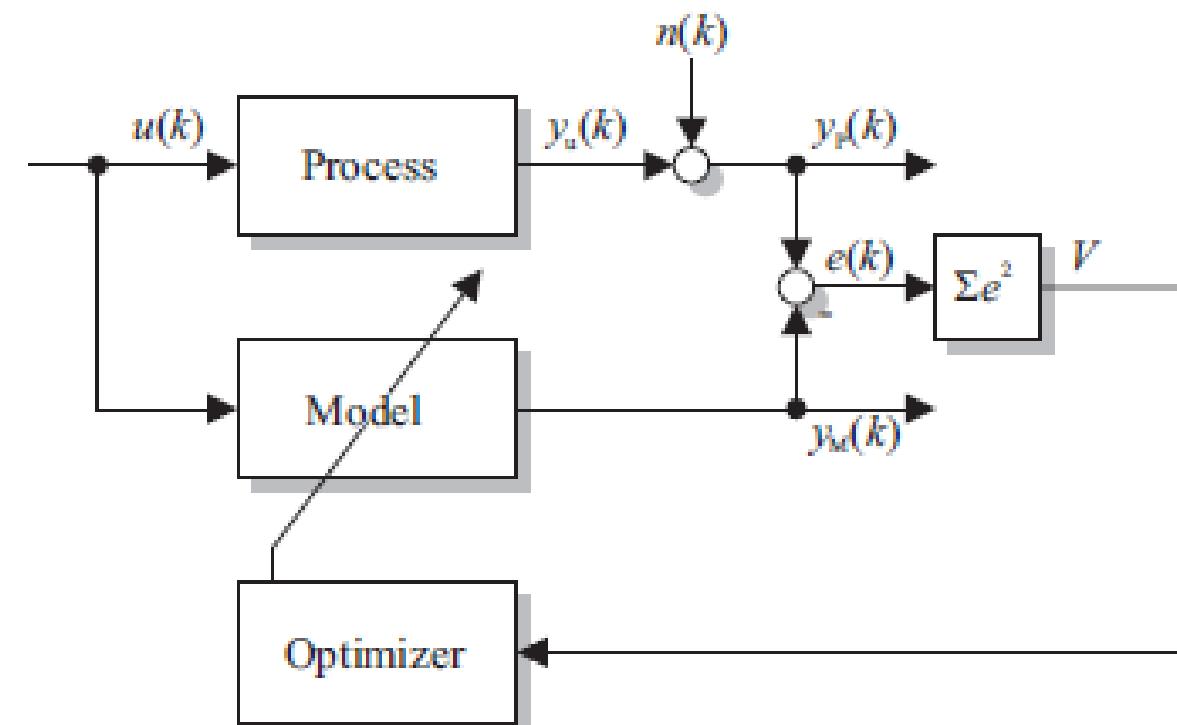
- Boscovich's (1711-1787) method and a minimization of the maximum deviation consisted of a minimization of the sum of absolute values of the residuals under the condition that the sum of the residuals should be equal to zero.
- An algorithm for finding the minimax residual was given in 1783 by Laplace, and in 1789 he simplified his earlier procedure.
- Another method was proposed by Euler and Lambert, according to which the estimates should be the quantities which minimize the absolute value of the largest deviation
- Legendre (1752-1833) published in 1805 a memoir, *Nouvelles méthodes pour la determination des cometes*, in which he introduced and named the method of least squares.
- Gauss (1777-1855) published in 1809 a book, *Theoria motus corporum coelestium in sectionibus conicis solem ambientium* (Gauss, 1809), where he discussed the method of least squares and, mentioning Legendre's work, stated that he himself had used the method since 1795
- Legendre was offended by Gauss's statement
- In the following years, Gauss tried to produce evidence for his claim but had only little success
- The astronomer Olbers included in a paper in 1816 a footnote asserting that Gauss had shown him the method of least squares in 1802. (Bessel published a similar note in a report in 1832.)

Metoda Najmniejszych Kwadratów – MNK

- Ogólny schemat stosowania metody
- Minimalizacja błędu obserwacji $e(k)$
- Stosujemy kryterium oceny w postaci sumy kwadratów błędu

$$V = e^2(1) + e^2(2) + \dots + e^2(N) = \sum_{k=1}^N (e(k))^2$$

- Dlaczego suma kwadratów?
 - Prostsza do minimalizacji
 - Przy założeniu rozkładu normalnego zakłócenia prowadzi do nieobciążonej estymaty parametrów
 - wada: nazbyt uwzględnia efekt pojedynczych dużych błędów w stosunku do małych ale częstych



Metoda Najmniejszych Kwadratów – MNK

- Rozwiązanie:

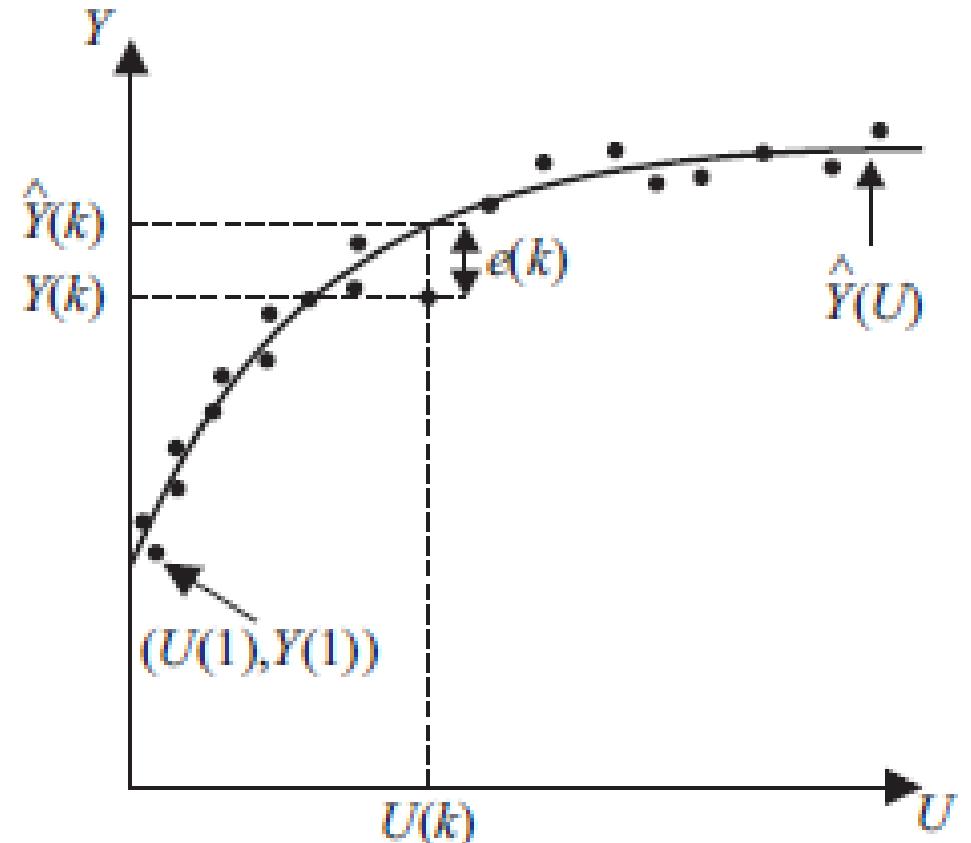
$$\mathbf{u}^T \mathbf{u} \hat{\mathbf{K}} = \mathbf{u}^T \mathbf{y}_P \Leftrightarrow \hat{\mathbf{K}} = (\mathbf{u}^T \mathbf{u})^{-1} \mathbf{u}^T \mathbf{y}_P$$

- Błąd:

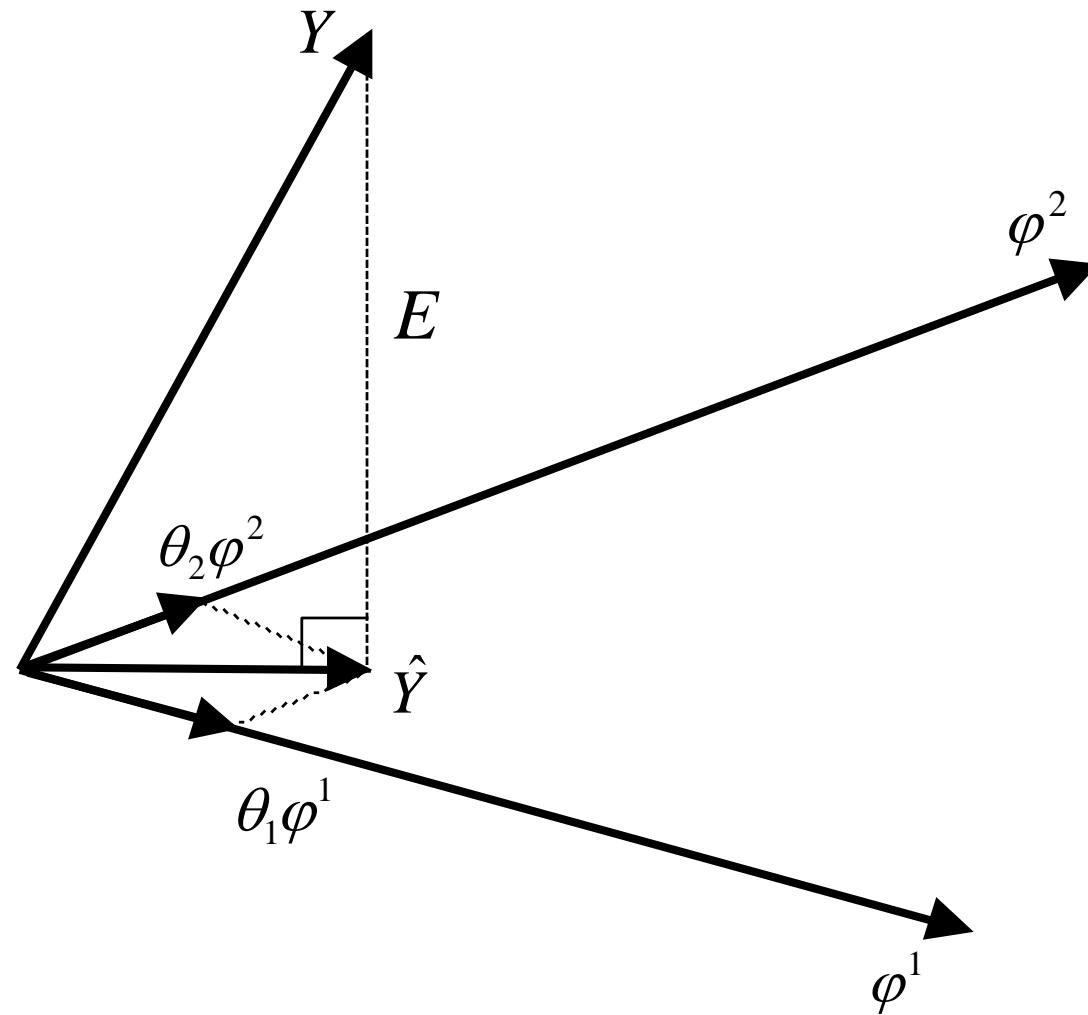
$$e(k) = Y_P(k) - Y_M(k) = y_P(k) - y_M(k) = y_P(k) - K_M(U(k) - U_{00})$$

- Komentarze:

- Sygnal wejściowy jest dokładnie mierzony jak i jego wartość średnia
- Obiekt jest **wystarczająco pobudzony**
- Zakłócenie $n(k)$ jest **stacjonarne**
- Wejście jest **nieskorelowane** z zakłóceniem
- Wartość oczekiwana albo wejścia, albo zakłócenia jest **zerowa**



MNK – interpretacja geometryczna



Metoda największej wiarygodności

- Metoda opiera się na warunkowym prawdopodobieństwie pomiaru $p_y(y|u, \theta)$ zwanym funkcja wiarygodności
- Idea polega na poszukiwaniu $\hat{\theta}$, która maksymalizuje tak zdefiniowaną wiarygodność

$$p_y(y|\theta)|_{\theta=\hat{\theta}} \rightarrow \max$$

- Maksimum znajdujemy w klasyczny sposób poprzez różniczkowanie

$$\frac{\partial p_y(y|\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}} = 0$$

- Teraz będziemy poszukiwać statycznej nieliniowości.

Metoda największej wiarygodności

- Pomiar wyjścia jest dany:
 $y = \Psi\Theta + n$
- Każdy próbka zakłócenia ma rozkład normalny opisany funkcją gęstości (przy czym zakładamy zerową wartość oczekiwana $\mu = 0$)

$$p(n(k)) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\frac{(n(k) - \mu)^2}{2\sigma_n^2}\right)$$

- Dla białego szumu poszczególne próbki są nieskorelowane na całej długości N , zatem funkcja gęstości dla całego ciągu n jest iloczynem gęstości poszczególnych próbek.

$$p(n) = \prod_{k=0}^{N-1} \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\frac{(n(k))^2}{2\sigma_n^2}\right) = \frac{1}{(2\pi)^{\frac{N}{2}} \sqrt{\det \Sigma}} e^{-\frac{1}{2} n^T \Sigma^{-1} n}$$

- Co daje N -wymiarową funkcję gęstości, gdzie Σ jest macierzą kowariancji.

$$\Sigma = \sigma_n^2 I \quad \det \Sigma = N \sigma_n^2$$

Metoda największej wiarygodności

- Zatem rozkład prawdopodobieństwa pomiaru ma postać

$$p(y|\theta) = \frac{1}{(2\pi)^{\frac{N}{2}} \sqrt{N} \sigma_n} \exp\left(-\frac{1}{2\sigma_n^2} (y - \Psi\theta)^T (y - \Psi\theta)\right)$$

- Teraz możemy wyznaczyć maksimum. Wiedząc że $p(x)$ i $\log p(x)$ mają maksimum dla tego samego x

$$\frac{\partial \log f(y|\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}} = \frac{1}{2\sigma_n^2} ((y - \Psi\hat{\theta})^T (y - \Psi\hat{\theta})) \stackrel{!}{=} 0$$

- Otrzymujemy:

$$\hat{\Theta} = (\Psi^T \Psi)^{-1} \Psi^T y$$

- Zatem estymator MNK i największej wiarygodności dają takie same rozwiązanie

Metoda największej wiarygodności

- Można wyznaczyć dolną granicę (Cramer-Rao) wariancji estymaty

$$E\{(\hat{\theta} - \theta)^2\} \geq \frac{1}{E\left\{\left(\frac{\partial}{\partial \theta} \log p_y(y|\theta)\right)^2\right\}}$$

$$E\{(\hat{\theta} - \theta)^2\} \geq \frac{-1}{E\left\{\frac{\partial^2}{\partial \theta^2} \log p_y(y|\theta)\right\}}.$$

- Estymator jest nazywany wydajnym, jeśli osiąga dolną granicę wariancji estymaty.
- BLUE (Best Linear Unbiased Estimator) osiąga minimalną wariancję wszystkich nieobciążonych estymatorów

Ograniczenia

- Równościowe

$$C\theta = d$$

- Może być spełnione i otrzymujemy rozwiązanie:

$$\hat{\theta} = \hat{\theta} - (\Psi^T \Psi)^{-1} C \left(C (\Psi^T \Psi)^{-1} C^T \right)^{-1} (C \hat{\theta} - d)$$

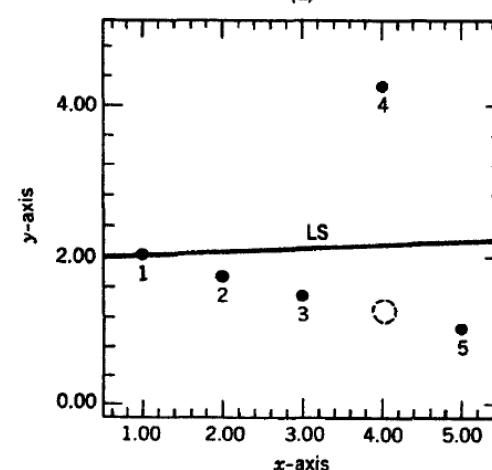
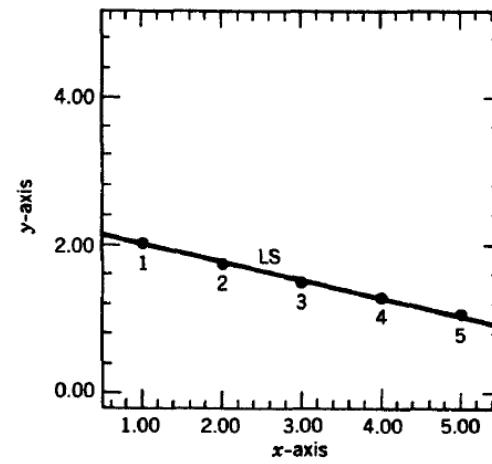
- Nierównościowe

$$A\theta \leq b$$

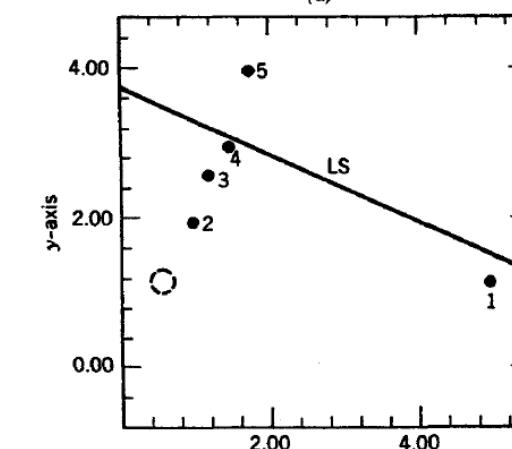
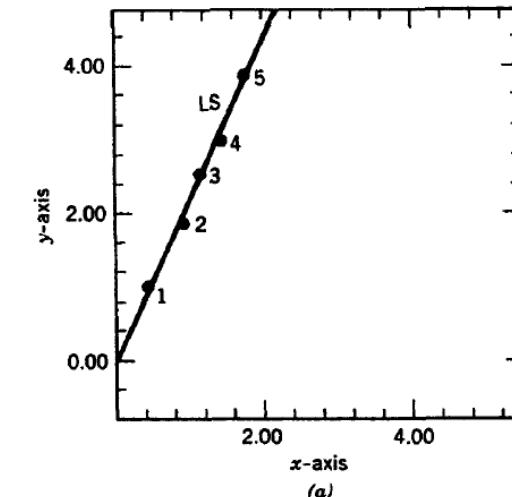
- Rozwiążujemy metodami optymalizacji nieliniowej

Outliers i ich wpływ na MNK

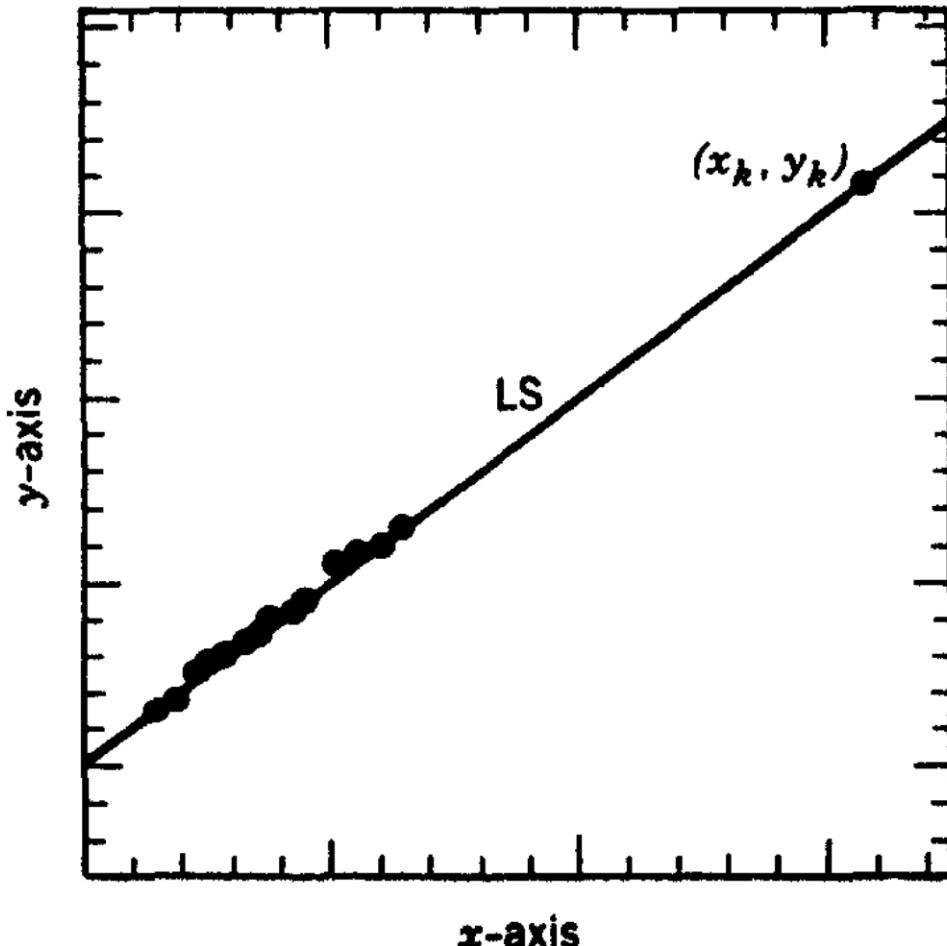
Outlier wg osi OY



Outlier wg osi OX



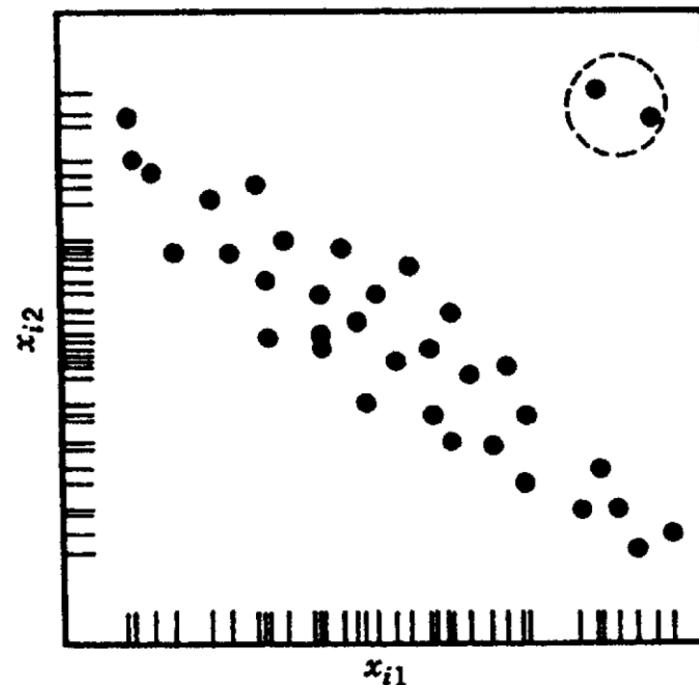
Outliers i ich wpływ na MNK



- Leverage point – punkt dźwigniowy
- Outlier jednocześnie w obu osiach
- Brak wpływu na regresję

Outliers i ich wpływ na MNK

- Dwa punkty *dźwigniowe*
- Nie są outlierem w żadnej z osi



Outliers

- Definicja:

an observation which deviates so much from other observations as to arouse suspicions that it was generated by a different mechanism

- Breakdown:

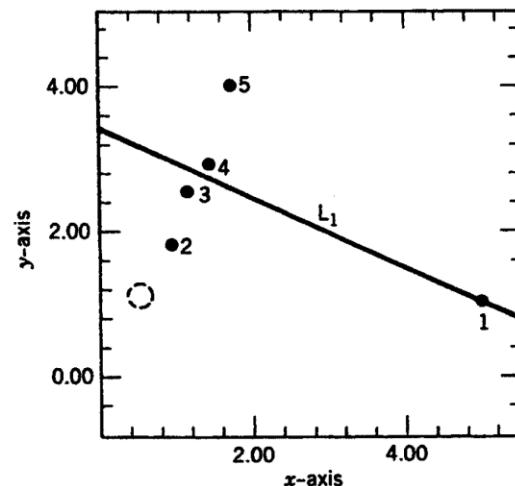
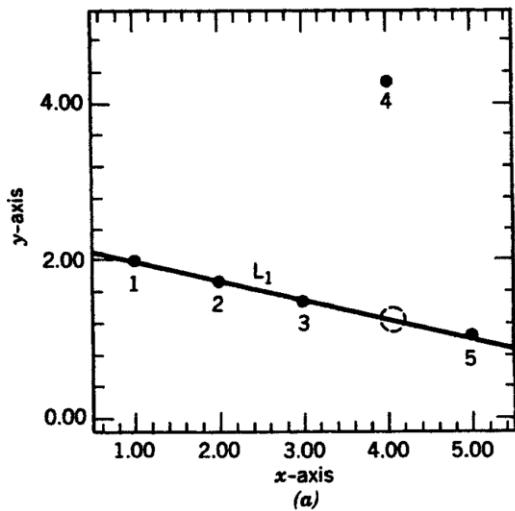
the smallest fraction of contamination that can cause the estimator T to take on values arbitrarily far from $T(2)$

- Dla MNK wynosi 0%

- Przyczyny:

- The data come from some heavy tailed distribution such as Student's t. There is no question that any observation is in any way erroneous.
- The data arise from two distributions. One of these, the 'basic distribution', generates 'good' observations, while another the 'contaminating distribution', generates 'contaminants'. If the contaminating distribution has tails which are heavier than those of the basic distribution, then there will be a tendency for the contaminants to be outliers - that is, to separate visibly from the good observations, which will then constitute the inliers.

Outliers i ich wpływ na MNK



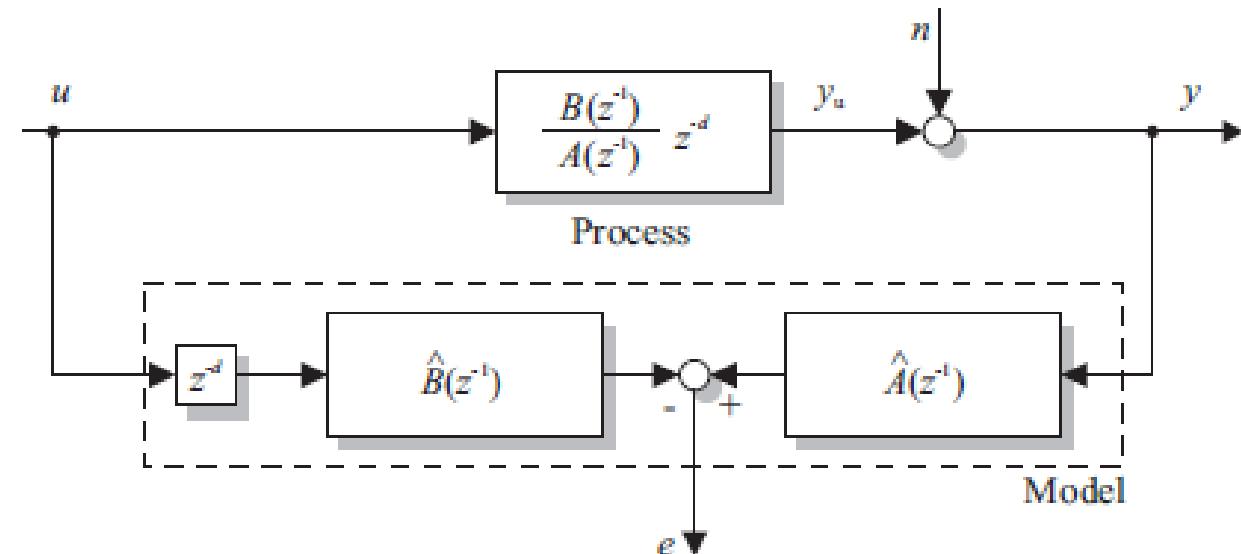
- Wskaźnik jakości Laplace'a, czyli
- IAE – błąd sumy wartości bezwzględnych
- Breakdown wynosi też 0%
- Estymatory odporne prowadzą do wartości breakdown sięgających 50%
 - LMS – Least Median of Squares
 - LTS - Least Trimmed Squares

MNK dla układów dynamicznych

- Model ARX

$$\widehat{\Theta} = \left[\frac{1}{n} \sum_{t=1}^n \varphi(t) \cdot \varphi^T(t) \right]^{-1} \cdot \left[\frac{1}{n} \sum_{t=1}^n \varphi(t) \cdot y(t) \right]$$

- Statystyczne własności estymaty są inne niż w przypadku statycznym, gdyż $\varphi(t)$ jest pewną realizacją procesu stochastycznego a nie ustaloną statyczną wielkością $\varphi(u)$



MNK: własności estymaty

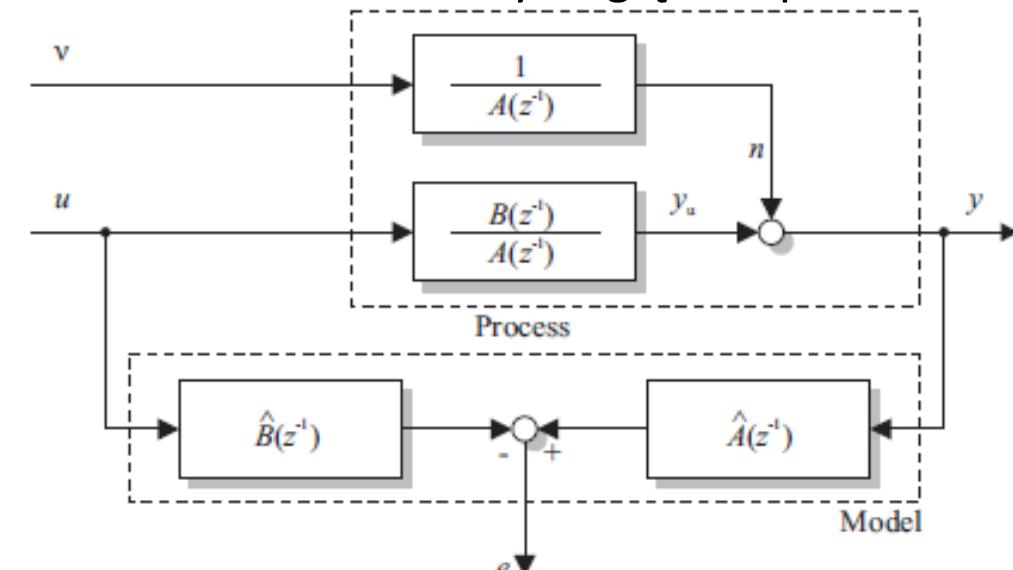
- Stąd $\widehat{\Theta} \xrightarrow{n \rightarrow \infty} \Theta_0$ jeśli

- $E\{\varphi(t) \cdot \varphi^T(t)\}$ nieosobliwa.

Na ogół spełnione, z wyjątkiem:

- Sygnał wejściowy nie jest dostatecznie dobrze pobudzony,
 - $v(t) = 0$ i rzad modelu za wysoki, tzn. $A_o(q^{-1})$ i $B_o(q^{-1})$ mają wspólnie człyony
 - $u(t)$ generowana przez liniowe sprzężenie zwrotne od wyjścia
- $E\{\varphi(t) \cdot v(t)\} = 0$
- Jedynie gdy $v(t)$ jest białym szumem.
 - Jeśli nie jest białym szumem, to jest skorelowana z poprzednimi wejściami, gdyż $y(t)$ zależy od $v(s)$ dla $s < t$ poprzez równanie procesu.
wtedy $E\{\varphi(t) \cdot v(t)\} \neq 0$

- Metoda łatwa w użyciu, ale daje estymator nieobciążony przy ostrych założeniach
 - Odchyłkę można czasem tolerować, np. gdy stosunek sygnału do szumu jest wysoki
 - Model może nie być zbyt dokładny (stosowany jako model wewnętrzny w układzie regulacji)
- Modyfikacje służące otrzymywaniu estymatorów **zgodnych**:
 - Modyfikacje układu równań normalnych, tzw. metody **zmiennej instrumentalnej IV** (ang. Instrumental Variable)
 - Przyjęcie innej, bardziej precyzyjnej postaci modelu. Niekoniecznie liniowy względem parametrów i stosujemy wtedy **metodę błędu predykcji**



Identyfikowalność parametrów

- Identyfikowalność jest uzależniona od:
 - Proces S
 - Zdefiniowanie eksperymentu X
 - Struktura modelu M
 - Metoda identyfikacji
- Różne definicje:
 - Funkcja celu ma jednoznaczne rozwiązanie (Bellmann i Astrom, 1970)
 - Spójność estymatora, tzn. jeśli estymaty zbiegają do rzeczywistych pomiarów

$$\widehat{\theta}(N) \xrightarrow{N \rightarrow \infty} \theta_0$$

$$\lim_{N \rightarrow \infty} E\{\widehat{\theta}(N)\} = \theta_0$$

$$\lim_{N \rightarrow \infty} \text{cov} \Delta \widehat{\theta}(N) = \mathbf{0}$$

- przy czym zbieżność określona jest w różnym sensie, np. prawdopodobieństwa, średnich kwadratów, itp..

Uwarunkowania

- Sygnał wejściowy musi być znany i mierzony tak jak i jego wartość średnia
- Sygnał wejściowy musi być wystarczająco pobudzony.
 - Dla sygnału wejściowego $u(k) = U(k) - \bar{U}$ powinien być spełniony warunek, że

$$\bar{U} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=m+d}^{m+d+N-1} U(k)$$

- oraz

$$R_{uu}(\tau) = \lim_{N \rightarrow \infty} (U(k) - \bar{U})(U(k - \tau) - \bar{U})$$

- Istnieją a macierz R_{22} jest dodatnio określona.
- Analogiczne warunki są dla metod korelacyjnych.
- Przykłady:
 - Biały szum,
 - Sygnał ruchomej średniej rzędu m
 - PRBS

Uwarunkowania

- Rzędy wielomianów oraz opóźnienie są znane
- Proces powinien spełniać następujące warunki:
 - Stabilny, czyli bieguny mianownika są w kole jednostkowym,
 - Nie wszystkie parametry b_i , $i=0,1,2,\dots,m$ są zerowe.
 - Licznik i mianownik nie ma wspólnych zer/biegunów.
- Jeśli minimalny rząd m jest znany, wtedy identyfikowalność jest zapewniona poprzez **stabilność, sterowalność** oraz **obserwonalność**.
- Zakłócenie losowe musi być stacjonarne
- Błąd nie może być skorelowany oraz jego wartość oczekiwana musi wynosić zero

Nieznana wartość stała wejścia

- Różnicowanie

$$U(k) - U(k-1) = u(k) - u(k-1) = \Delta u(k)$$

$$Y(k) - Y(k-1) = y(k) - y(k-1) = \Delta y(k)$$

- Uśrednianie

$$\hat{Y}_{00} = \frac{1}{N} \sum_{k=0}^{N-1} Y(k)$$

uśrednianie rekurencyjne → $\hat{Y}_{00} = \hat{Y}_{00}(k-1) + \frac{1}{k} (Y(k) - \hat{Y}_{00}(k-1))$

uśrednianie rekurencyjne z zapominaniem → $\hat{Y}_{00} = \lambda \hat{Y}_{00}(k-1) + (1-\lambda) Y(k)$

- Włączenie wartości średniej do identyfikowanego wektora parametrów

$$Y(k) = -a_1 Y(k-1) - \dots - a_m Y(k-m) + b_1 U(k-d-1)$$

$$+ \dots + b_m U(k-d-m) + C,$$

$$C = (1 + a_1 + \dots + a_m) Y_{00} - (b_1 + \dots + b_m) U_{00}$$



MI

Metody Identyfikacji

wykład #7

1. *Identyfikacja modeli parametrycznych cd.*
 - a) *Metody dwuetapowe*
 - b) *Metody rekurencyjne*
 - c) *Rozszerzenia metody najmniejszych kwadratów*

Identyfikacja dwuetapowa z pośrednim modelem nieparametrycznym

- Jeśli struktura jest nieznana (a tak jest najczęściej) to można najpierw znaleźć model nieparametryczny (nie ma potrzeby na wiedzę *a priori*) a potem dopiero znajdujemy strukturę i parametry.
- Dwa podejścia:
 1. Odpowiedź na pobudzenie nieokresowe a potem MNK
 2. Metoda korelacyjna a potem MNK

Odpowiedź na pobudzenie nieokresowe a potem MNK

- Kilka M odpowiedzi na to samo pobudzenie $u_j(k)$ a potem uśrednienie wyjść $y_j(k)$ w celu eliminacji wpływu addytywnego zakłócenia stochastycznego $n_j(k)$

$$\bar{y}(k) = \frac{1}{M} \sum_{j=1}^M y_j(k)$$

- Otrzymujemy wartość oczekiwana

$$E\{\bar{y}(k)\} = y_u(k) + E\{\bar{n}(k)\}$$

- Jeśli wartość oczekiwana zakłócenia jest zerowa, to wtedy obie wartości oczekiwane są identyczne.
- Model ARX

$$\begin{aligned} y(k) = & -a_1 y(k-1) - a_2 y(k-2) - \dots - a_m y(k-m) \\ & + b_1 u(k-d-1) + b_2 u(k-d-2) + \dots + b_m u(k-d-m) \end{aligned}$$



Odpowiedź na pobudzenie nieokresowe a potem MNK

- Przepisujemy model do postaci wektorowej dla $1 \leq k \leq l$

$$\begin{pmatrix} y(1) \\ y(2) \\ y(3) \\ \vdots \\ y(l) \end{pmatrix} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ -y(1) & 0 & \dots & 0 \\ -y(2) & -y(1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -y(l-1) & -y(l-2) & \dots & -y(l-m) \end{pmatrix} \begin{pmatrix} u(-d) & \dots & 0 \\ u(1-d) & \dots & 0 \\ u(2-d) & \dots & 0 \\ \vdots & \ddots & \vdots \\ u(l-d-1) & \dots & u(l-d-m) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \\ b_1 \\ \vdots \\ b_m \end{pmatrix}$$

- Co zapisujemy jako $y = R\Theta$ i otrzymujemy:

$$e = y - R\Theta$$

$$V = e^T e$$

$$\hat{\Theta} = (R^T R)^{-1} R^T y$$

Odpowiedź na pobudzenie nieokresowe a potem MNK

- Estymata jest spójna, tzn. dla liczby eksperymentów $M \rightarrow \infty$ otrzymujemy:
 - Wartość oczekiwana błędu równą zero
 - Wartość oczekiwana parametrów równą rzeczywistym parametrom
 - Wartość oczekiwana wariancji parametrów równą zero
- Zalety metody:
 - Unikanie zakłóceń i szumów poprzez uśrednianie
- Wady:
 - Konieczność wielokrotnego przeprowadzenia eksperymentu

Metoda korelacyjna a potem MNK

- Jeśli na wejściu podajemy sygnał pseudolosowy to funkcja autokorelacji ma postać

$$R_{uu}(\tau) = \lim_{N \rightarrow \infty} \frac{1}{N+1} \sum_{k=0}^N u(k-\tau)u(k)$$

- A funkcja korelacji wzajemnej

$$R_{uy}(\tau) = \lim_{N \rightarrow \infty} \frac{1}{N+1} \sum_{k=0}^N u(k-\tau)y(k)$$

- Funkcje korelacji mogą być wyznaczone rekurencyjnie

$$\hat{R}_{uy}(\tau, k) = \hat{R}_{uy}(\tau, k-1) + \frac{1}{k+1} (u(k-\tau)y(k) - \hat{R}_{uy}(\tau, k-1))$$

- Model ARX jak poprzednio

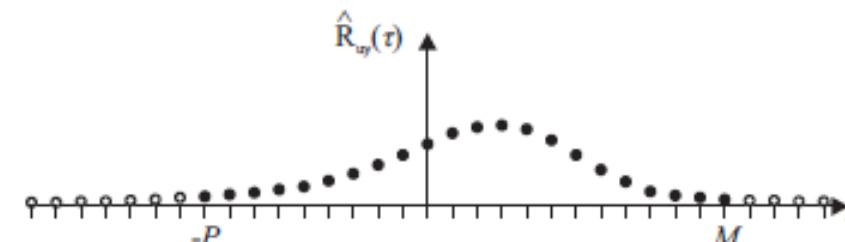
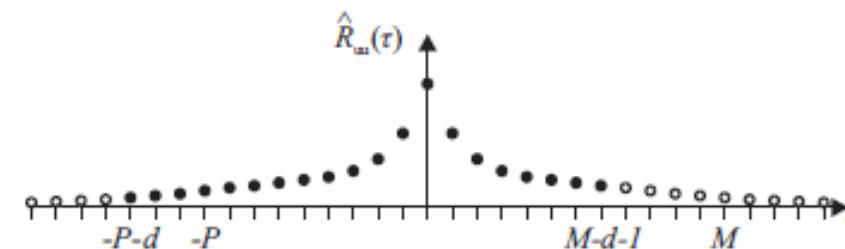
$$y(k) = -a_1y(k-1) - a_2y(k-2) - \dots - a_my(k-m) \\ + b_1u(k-d-1) + b_2u(k-d-2) + \dots + b_mu(k-d-m)$$

Metoda koreacyjna a potem MNK

- Tym samym dla funkcji korelacji wzajemnej otrzymujemy

$$\begin{aligned} R_{uy}(\tau) = & -a_1 R_{uy}(\tau - 1) - a_2 R_{uy}(\tau - 2) - \dots - a_m R_{uy}(\tau - m) \\ & + b_1 R_{uu}(\tau - d - 1) + b_2 R_{uu}(\tau - d - 2) + \dots + b_m R_{uu}(\tau - d - m) \end{aligned}$$

- Równanie to stanowi podstawę metody identyfikacji
- Przymajemy, że wartości funkcji korelacji wzajemnej wykorzystywane do estymaty parametrów modelu są różne od zera dla $-P \leq \tau \leq M$ i równe \approx zeru dla $\tau < -P$ i $\tau > M$



Metoda korelacyjna a potem MNK

- Ostatecznie otrzymujemy układ równań

$$\begin{pmatrix} R_{uy}(-P+m) \\ \vdots \\ R_{uy}(-1) \\ R_{uy}(0) \\ R_{uy}(1) \\ \vdots \\ R_{uy}(M) \end{pmatrix} = \begin{pmatrix} -R_{uy}(-P+m+1) & \dots & -R_{uy}(-P) & R_{uu}(-P+m-d-1) & \dots \\ \vdots & & \vdots & \vdots & \dots \\ -R_{uy}(-2) & \dots & -R_{uy}(-1-m) & R_{uu}(-2-d) & \dots \\ -R_{uy}(-1) & \dots & -R_{uy}(-m) & R_{uu}(-d-1) & \dots \\ -R_{uy}(0) & \dots & -R_{uy}(1-m) & R_{uu}(-d) & \dots \\ \vdots & & \vdots & \vdots & \dots \\ -R_{uy}(M-1) & \dots & -R_{uy}(M-m) & R_{uu}(M-d-1) & \dots \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_m \\ b_1 \\ \vdots \end{pmatrix}$$

- Co zapisujemy jako $R_{uy} = S\Theta$ a zastosowanie MNK prowadzi do estymaty parametrów

$$\hat{\Theta} = (S^T S)^{-1} S^T R_{uy}$$

Metoda koreacyjna a potem MNK

- Wersja nierekurencyjna

1. Wczytanie wektorów wejść $u(k)$ i wyjść $y(k)$
2. Wyznaczenie autokorelacji wejścia $R_{uu}(\tau)$ oraz korelacji wzajemnej wejście-wyjście $R_{uy}(\tau)$
3. Wyznaczenie estymaty

- Wersja rekurencyjna

1. W k -tym kroku wyznaczamy autokorelację wejścia $R_{uu}(\tau, k)$ oraz korelację wzajemną wejście-wyjście $R_{uy}(\tau, k)$
2. Wyznaczamy estymatę parametrów w każdym kroku, lub też w dłuższych okresach.

$$\widehat{\boldsymbol{\theta}} = (\mathbf{S}^T \mathbf{S})^{-1} \mathbf{S}^T \mathbf{R}_{uy}$$

Rekurencyjna MNK (RLS – Recursive Least Squares)

- Pierwszy opis metody: *Gauss* (1809)
- Pierwsze zastosowanie: *Lee* (1964)
- Metoda nierekurencyjna (dla kroku k):

$$\hat{\theta}(k) = P(k)\Psi^T(k)y(k)$$

$$P(k) = (\Psi^T(k)\Psi(k))^{-1}$$

$$y(k) = \begin{pmatrix} y(1) \\ y(2) \\ \vdots \\ y(k) \end{pmatrix}$$

$$\Psi(k) = \begin{pmatrix} \psi^T(1) \\ \psi^T(2) \\ \vdots \\ \psi^T(k) \end{pmatrix}$$

$$\psi^T = (-y(k-1) - y(k-2) \dots - y(k-m) | u(k-d-1) \dots u(k-d-m))$$

Rekurencyjna MNK (RLS – Recursive Least Squares)

- Dla kroku $k+1$ otrzymujemy:

$$\hat{\theta}(k+1) = P(k+1)\Psi^T(k+1)y(k+1)$$

- Równanie to można podzielić na:

$$\begin{aligned}\hat{\theta}(k+1) &= P(k+1) \left(\begin{matrix} \Psi(k) \\ \Psi^T(k+1) \end{matrix} \right)^T \left(\begin{matrix} y(k) \\ y(k+1) \end{matrix} \right) \\ &= P(k+1)(\Psi(k)y(k) + \Psi^T(k+1)y(k+1))\end{aligned}$$

- Podstawiając estymatę z kroku k w postaci $\Psi(k)y(k) = P^{-1}(k)\hat{\theta}(k)$ otrzymamy

$$\hat{\theta}(k+1) = \hat{\theta}(k) + (P(k+1)P^{-1}(k) - I)\hat{\theta}(k) + P(k+1)\Psi^T(k+1)y(k+1)$$

$$P^{-1}(k) = P^{-1}(k+1) - \Psi(k+1)\Psi^T(k+1)$$

$$\begin{aligned}P(k+1) &= \left(\left(\begin{matrix} \Psi(k) \\ \Psi^T(k+1) \end{matrix} \right)^T \left(\begin{matrix} \Psi(k) \\ \Psi^T(k+1) \end{matrix} \right) \right)^{-1} \\ &= (P^{-1}(k) + \Psi(k+1)\Psi^T(k+1))^{-1}\end{aligned}$$

Rekurencyjna MNK

- Otrzymujemy:

$$\widehat{\Theta}(k+1) = \widehat{\Theta}(k) + P(k+1)\psi(k+1)$$

Nowa estymata
parametru

Stara estymata
parametru

Wektor
korekcji

$$(y(k+1) \quad \psi^T(k+1)\widehat{\theta}(k))$$

Nowy
pomiar

Preidykcja pomiaru na podstawie
ostatniej dostępnej estymaty
 $\hat{y}(k+1|k)$

- A ostatecznie

$$\widehat{\Theta}(k+1) = \widehat{\Theta}(k) + P(k+1)\psi(k+1)e(k+1)$$

Rekurencyjna MNK

- Można zrezygnować z odwracania macierzy $P(k+1)$:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \gamma(k)(y(k+1) - \psi^T(k+1)\hat{\theta}(k))$$

$$\gamma(k) = P(k+1)\psi(k+1) = \frac{1}{\psi^T(k+1)P(k)\psi(k+1) + 1} P(k)\psi(k+1)$$

- Macierz $P(k+1)$ jest przeskalowaną estymatą macierzy kowariancji błędu estymacji

$$P(k+1) = (I - \gamma(k)\psi^T(k+1))P(k)$$

$$E\{P(k+1)\} = \frac{1}{\sigma_e^2} \text{cov } \Delta\theta(k+1)$$

Rekurencyjna MNK

- Inicjalizacja metody:

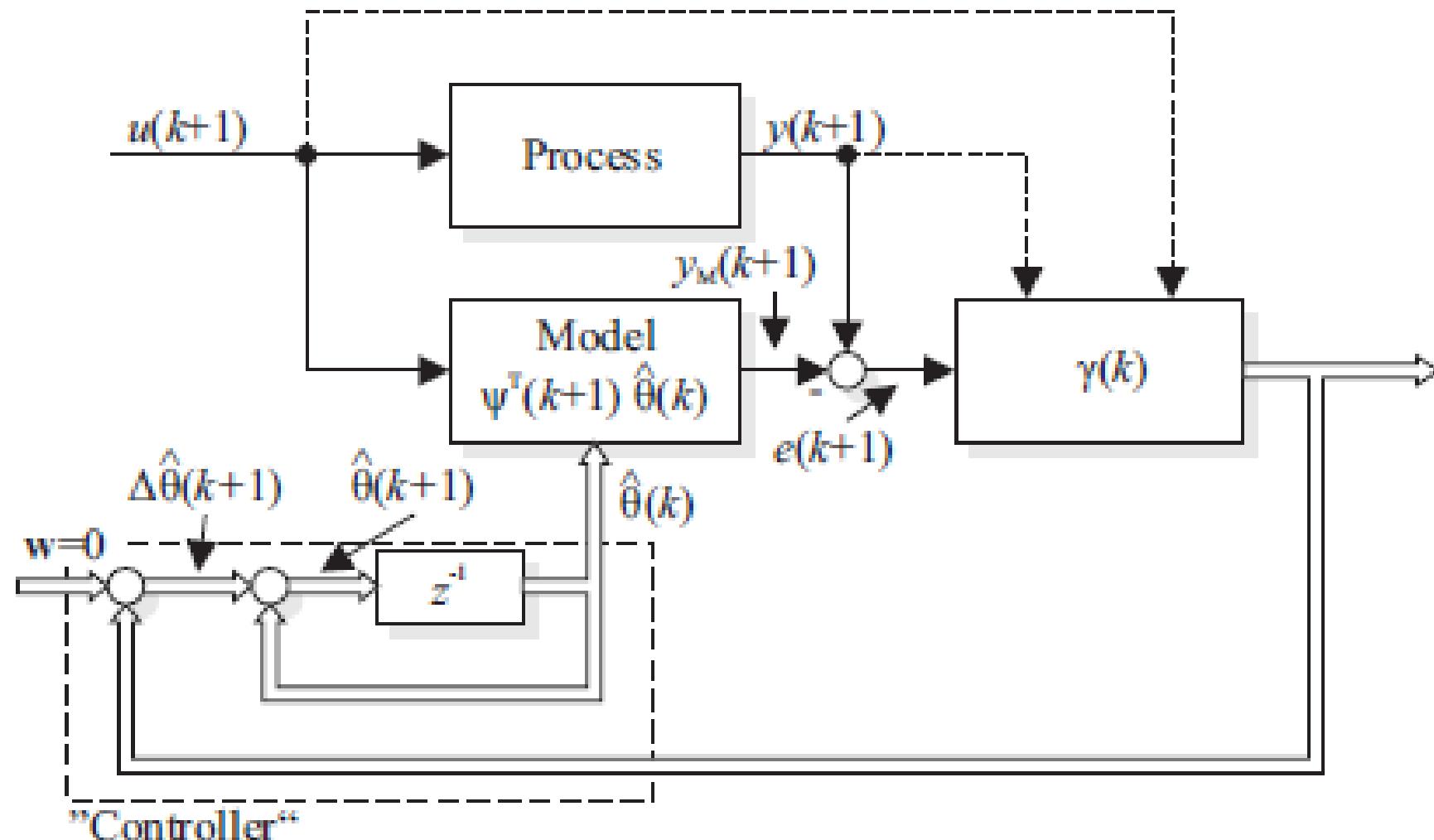
- Rozpocząć od metody nierekurencyjnej dla przynajmniej $2m$ rekordów i potem wystartować rekurencję
- Wykorzystać estymatę a priori parametrów, ich kowariancji oraz kowariancji błędu
- wystartowanie z warunkiem początkowym $P(0) = P_0$, $\theta(0) = \theta_0$, gdzie P_0 jest macierzą dodatnio określona i dobrana tak aby

$$P^{-1}(k) = P^{-1}(0) + \Psi^T(k)\Psi(k)$$

- np. $P_0 = \alpha I$, gdzie α jest dużą liczbą.

$$\alpha \gg \frac{1}{u^2(0)}$$

Schemat blokowy rekurencyjnej MNK



Metoda ważona MNK (WLS – Weighted Least Squares)

- Jak do tej pory każdy element $e(k)$ miał takie same znaczenie
- Ogólnie można wprowadzić wagi różne dla poszczególnych chwil czasowych
- Wprowadzamy funkcję celu:

$$V = w(m+d)e^2(m+d) + w(m+d+1)e^2(m+d+1) + \dots \\ + w(m+d+N)e^2(m+d+N)$$

- A w ogólnej postaci

$$V = e^T W e$$

- Gdzie W jest symetryczną dodatnio określona macierzą
 - Tylko część symetryczna ma wpływ
 - Dodatnio określona dla istnienia jednoznacznego wyniku

$$W = \begin{pmatrix} w(m+d) & 0 & \dots & 0 \\ 0 & w(m+d+1) & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & w(m+d+N) \end{pmatrix}$$

Metoda ważona MNK

- W prosty sposób otrzymujemy

$$\widehat{\Theta} = (\Psi^T W \Psi)^{-1} \Psi^T W y$$

- Analogicznie otrzymujemy zależności dla ważonej metody najmniejszych kwadratów

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \gamma(k)(y(k+1) - \psi^T(k+1)\hat{\theta}(k))$$

$$\gamma(k) = \frac{1}{\psi^T(k+1)P_W(k)\psi(k+1) + \frac{1}{w(k+1)}} P_W(k)\psi(k+1)$$

$$P_W(k+1) = (I - \gamma(k)\psi^T(k+1))P_W(k).$$

- Porównując z metodą podstawową zmienia się parametr korekcyjny $\gamma(k)$ a tym samym $P_w(k+1)$

Metoda rekurencyjna MNK z zapominaniem eksponentjalnym

- Wprowadzenie parametrów zapominania zależnych od czasu, np.

$$w(k) = \lambda^{(m+d+N)-k} = \lambda^{N'-k}, \quad 0 < \lambda < 1$$

- Gdzie λ jest nazywana współczynnikiem zapominania.
- Poniżej pokazana są przykładowe parametry dla $N=50$.

k	1	10	20	30	40	47	48	49	50
$\lambda = 0.99$	0.61	0.67	0.73	0.82	0.90	0.97	0.98	0.99	1
$\lambda = 0.95$	0.08	0.13	0.21	0.35	0.60	0.85	0.90	0.95	1

$$W(m + d + n) = \begin{pmatrix} \lambda^N & & & & \\ & \lambda^{N-1} & & & \\ & & \ddots & & \\ & & & \lambda^2 & \\ & & & & \lambda \\ & & & & & 1 \end{pmatrix}$$

Metoda rekurencyjna MNK z zapominaniem eksponentjalnym

- Po nadejściu kolejnego pomiaru macierz jest uaktualniana

$$W(k+1) = \begin{pmatrix} \lambda W(k) & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}$$

- Ostatecznie metoda sprowadzona jest do postaci:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \gamma_W(k)(y(k+1) - \psi^T(k+1)\hat{\theta}(k))$$

$$\gamma_W(k) = \frac{1}{\psi^T(k+1)P_W(k)\psi(k+1) + \lambda} P_W(k)\psi(k+1)$$

$$P_W(k+1) = (I - \gamma_W(k)\psi^T(k+1))P_W(k)\frac{1}{\lambda}$$

- Dobór współczynnika zapominania jest kompromisem pomiędzy tłumieniem zakłóceń $\gamma \rightarrow 1$ a lepszym śledzeniem procesów niestacjonarnym $\gamma < 1$
- W praktyce stosuje się wielkości około $0.9 < \gamma < 0.995$

Ogólniejsze metody

- Model Box-Jenkins umożliwia bardziej szerokie uwzględnianie zakłóceń

$$y(k) = \frac{B(z^{-1})}{A(z^{-1})} z^{-d} u(k) + \frac{D(z^{-1})}{C(z^{-1})} v(k)$$

$$n(k) = \frac{D(z^{-1})}{C(z^{-1})} v(k)$$

- Metoda działa jedynie wtedy, gdy błąd estymaty $e(k)$ jest nieskorelowany. Dzieje się tak jedynie wtedy, gdy zakłócenie $n(k)$ jest generowane poprzez koloryzowanie białego szumu $v(k)$ – rzadko spełnione.
- Poniżej będą przedstawione metody posiadające szerszy zakres stosowalności

Ogólna MNK (GLS – Generalized Least Squares)

- Zakładamy, że nieskorelowany sygnał błędu

$$A(z^{-1})y(z) - B(z^{-1})z^{-d}u(z) = e(z)$$

- Jest zastąpiony przez sygnał skorelowany, szum kolorowy $\xi(k)$, generowany jak poniżej:

$$\xi(z) = \frac{1}{F(z^{-1})} e'(z), \text{ gdzie } e'(z) \text{ jest nieskorelowany.}$$

- Wielomian $F(z^{-1})$ jest nieznany, zatem zaproponowany jest algorytm iteracyjny.

Ogólna MNK (GLS – Generalized Least Squares)

Krok 1

- Stosujemy metodę MNK do pomiarów na przedziale

$$m + d \leq k \leq m + d + N$$

- dla modelu

$$A(z^{-1})y(z) - B(z^{-1})z^{-d}u(z) = \xi(z)$$

- Estymaty $\hat{\theta}_1$ są obciążone, a $\xi(k)$ skorelowany.

Ogólna MNK (GLS – Generalized Least Squares)

Krok 2

- Sygnał $\xi(k)$ jest wyznaczany dla parametrów $\hat{\theta}_1$. Wykorzystanie modelu AR

$$\xi(k) = \psi_{\xi}^T(k)f + e'(k)$$

- Prowadzi do

$$\psi_{\xi}^T(k) = [-\xi(k-1), -\xi(k-2), \dots, -\xi(k-\nu)]$$

$$f^T = [f_1, f_2, \dots, f_\nu]$$

- Rząd wielomianu powinien być wyznaczony dokładnie, tj. $\nu = m$.
- Parametry otrzymujemy z metody najmniejszych kwadratów:

$$\hat{f} = [\Xi^T \Xi]^{-1} \Xi^T \xi,$$

gdzie Ξ jest utworzony z wierszy $\psi_{\xi}^T(k)$.



Ogólna MNK (GLS – Generalized Least Squares)

Krok 3

- Wejście $u(k)$ i wyjście $y(k)$ filtrujemy poprzez

$$G_F(z^{-1}) = \hat{F}(z^{-1})$$

- zatem

$$\tilde{u}(z) = G_F(z^{-1})u(z) \text{ oraz } \tilde{y}(z) = G_F(z^{-1})y(z)$$

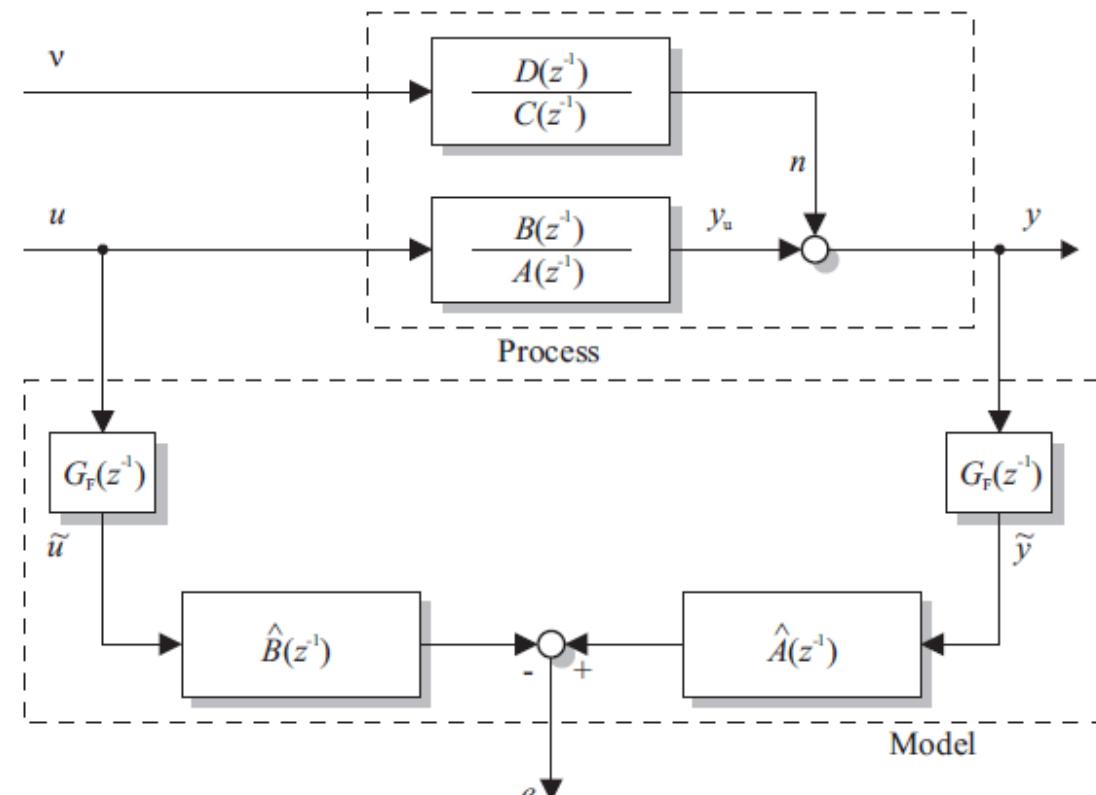
Ogólna MNK (GLS – Generalized Least Squares)

Krok 4

- Metoda najmniejszych kwadratów jest zastosowana do sygnałów filtrowanych

$$A(z^{-1})\tilde{y}(z) - B(z^{-1})z^{-d}\tilde{u}(z) = \xi(z)$$

- i otrzymujemy wektor parametrów $\hat{\theta}_2$.



Ogólna MNK (GLS – Generalized Least Squares)

Krok 5

- Kroki od 2 do 4 powtarzamy aż $\hat{\theta}_j$ nie zmieniają się w sposób znaczący pomiędzy iteracjami.
- Metoda ogólna MNK prowadzi do nieobciążonej estymaty szumu generowany poprzez:

$$G_v(z) = \frac{n(z)}{v(z)} = \frac{D(z^{-1})}{C(z^{-1})} = \frac{1}{A(z^{-1})F(z^{-1})}$$

Ogólna Rekurencyjna MNK (RGLS – Recursive Method of Generalized Least Squares)

- Z metody ogólnej MNK można wyprowadzić wersje rekursywną

$$\hat{\theta}(k+1) = \hat{\theta}(k) + (\psi^T(k+1) \tilde{P}(k) \bar{\psi}(k+1) + 1)^{-1}$$

$$\tilde{P}(k) \bar{\psi}(k+1) (\bar{y}(k+1) - \psi^T(k+1) \hat{\theta}(k))$$

$$\begin{aligned} \tilde{P}(k+1) = & \tilde{P}(k) \left(I - \bar{\psi}^T(k+1) \bar{\psi}(k+1) \tilde{P}(k) \right. \\ & \left. (\bar{\psi}^T(k+1) \tilde{P}(k) \bar{\psi}(k+1) + 1)^{-1} \right) \end{aligned}$$

$$\begin{aligned} \tilde{f}(k+1) = & \tilde{f}(k) + (\psi_{\xi}^T(k+1) \underline{Q}(k) \psi_{\xi}(k+1) + 1)^{-1} \\ & \underline{Q}(k) \psi_{\xi}(k+1) (\xi(k+1) - \psi_{\xi}^T(k+1) \hat{f}(k)) \end{aligned}$$

$$\begin{aligned} \underline{Q}(k+1) = & \underline{Q}(k) \left(I - \psi_{\xi}^T(k+1) \psi_{\xi}(k+1) \underline{Q}(k) \right. \\ & \left. (\psi_{\xi}^T(k+1) \underline{Q}(k) \psi_{\xi}(k+1) + 1)^{-1} \right). \end{aligned}$$

Rozszerzona MNK (ELS – Extended Least Squares)

- Zamiast MNK dla modelu ze skorelowanym sygnałem $\varepsilon(z)$

$$A(z^{-1})\tilde{y}(z) - B(z^{-1})z^{-d}\tilde{u}(z) = \varepsilon(z)$$

- Stosujemy model ARMAX ze skorelowanym sygnałem $\varepsilon(z) = D(z^{-1})e'(z^{-1})$.
- W takim przypadku możemy połączyć rekurencyjne metody dla obiektu dynamicznego oraz procesu stochastycznego.
- Dla modelu

$$y(k) = \psi^T(k)\hat{\theta}(k-1) + e(k)$$

Wprowadzamy rozszerzony wektor

$$\begin{aligned} \psi^T(k) &= (-y(k-1) \dots -y(k-m) | u(k-d-1) \dots \\ &\quad u(k-d-m) | \hat{v}(k-1) \dots \hat{v}(k-p)) \\ \hat{\theta}^T &= (\hat{a}_1 \dots a_m | \hat{b}_1 \dots \hat{b}_m | \hat{d}_1 \dots \hat{d}_p) . \end{aligned}$$

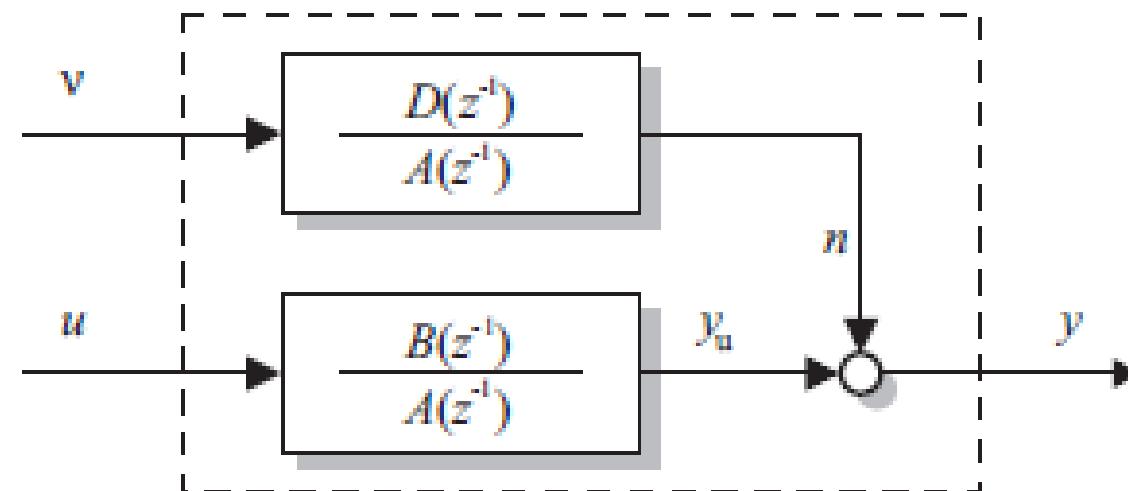
Rozszerzona MNK (ELS – Extended Least Squares)

- I sprowadzamy rozwiązanie do metody rekurencyjnej

$$\hat{\theta}(k+1) = \hat{\theta}(k) + \gamma(k)(y(k+1) - \psi^T(k+1)\hat{\theta}(k))$$

$$\gamma(k) = P(k+1)\psi(k+1) = \frac{1}{\psi^T(k+1)P(k)\psi(k+1) + 1}P(k)\psi(k+1)$$

$$P(k+1) = (I - \gamma(k)\psi^T(k+1))P(k)$$



Method of Bias Correction (CLS)

- Dotychczasowe metody starały się uniknąć obciążenia metody zakładając odpowiednie założenia na tworzenie sygnału zakłócającego.
- Alternatywne podejście próbuje wyznaczyć obciążenie metody (*bias*) a potem jego wykorzystanie do korekcji.
- Niemniej założenia też są bardzo ostre, gdyż sygnał $v(t)$ musi być białym szumem.

Method of Bias Correction (CLS)

- Zakładamy model w postaci

$$y(z) = \frac{B(z^{-1})}{A(z^{-1})} z^{-d} u(z) + n(z)$$

gdzie $n(z)$ jest białym szumem.

- Bias opisany jest przez

$$E\{b(N+1)\} = -E\{R^{-1}(N+1)\} \underbrace{\begin{pmatrix} I & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}}_S \theta_0 \sigma_n^2$$

- Jest on wykorzystany do korekcji estymaty wektora parametrów

$$\begin{aligned}\hat{\theta}_{\text{CLS}}(N+1) &= \hat{\theta}_{\text{LS}}(N+1) - b(N+1) \\ &= \hat{R}^{-1}(N+1) \frac{1}{N+1} \Psi^T(N+1) y(N+1) \\ &\quad + \hat{R}^{-1}(N+1) S \hat{\theta}_{\text{CLS}}(N+1) \sigma_n^2\end{aligned}$$

Method of Bias Correction (CLS)

- Otrzymujemy postać algorytmu

$$\hat{\theta}_{\text{CLS}}(N+1) = (R(N-1) - S\sigma_n^2)^{-1} \frac{1}{N+1} \Psi^T(N+1)y(N+1)$$

- Gdzie wariancja jest opisana poprzez i może być liczona koreacyjnie



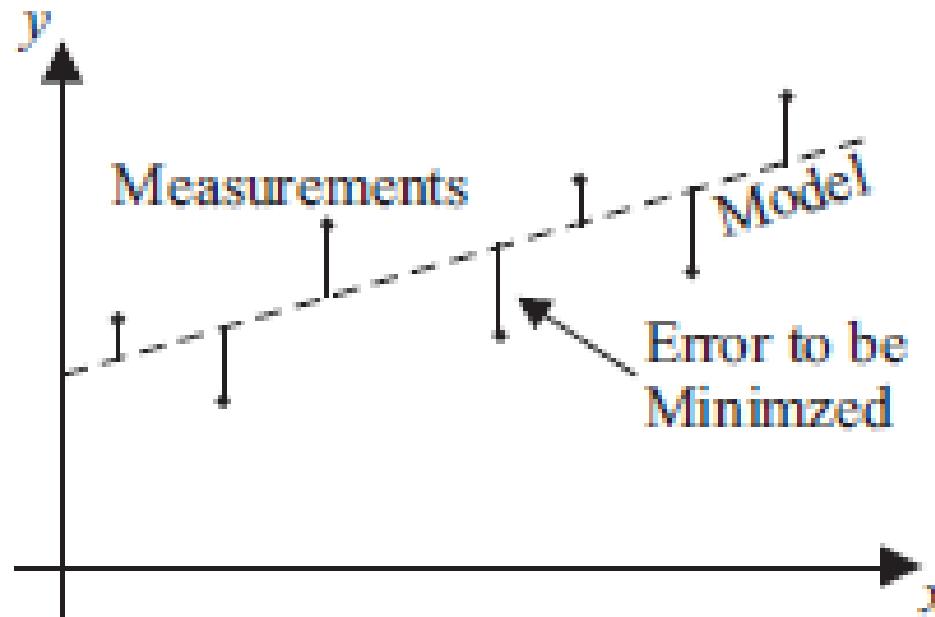
$$\sigma_n^2(N+1) = E\{n^2(k)\} = \frac{1}{N+1-2m} n^T(N+1)n(N+1)$$

Total Least Squares (TLS)

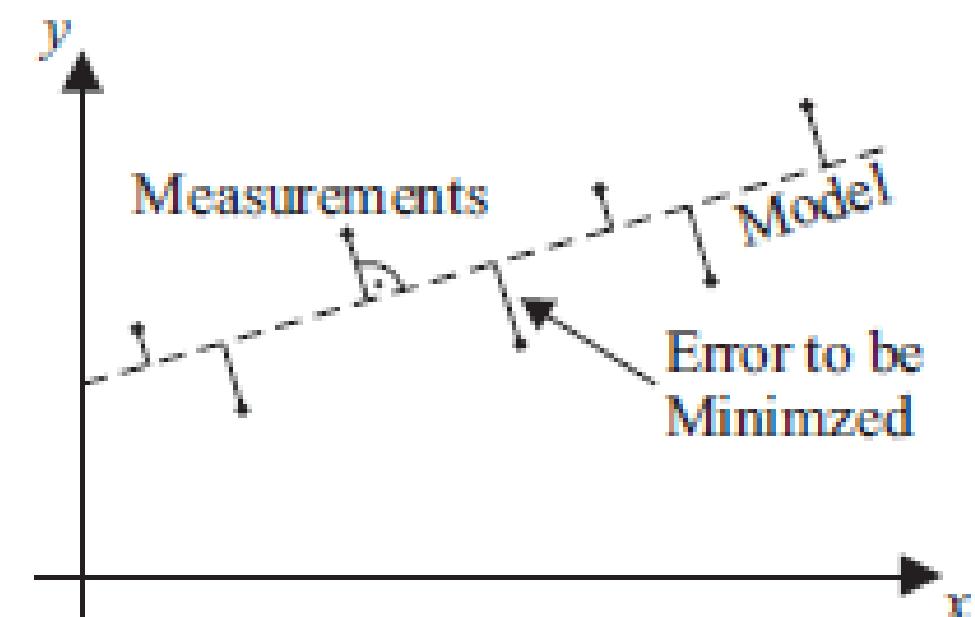
- Klasyczny model zakłada, że tylko wyjście jest zakłócone, tzn.

$$y - e = \Psi \hat{\Theta}$$

$$\hat{\Theta} = \arg \min \|e\|_2^2$$



- Obecnie zakładamy model w postaci $y - e = (\Psi + F)\hat{\Theta}$
- Czyli błąd dotyczy nie tylko wyjścia ale również pomiarów w wektorze regresji Ψ



Total Least Squares (TLS)

- Model możemy przepisać do postaci

$$\underbrace{\begin{pmatrix} (\Psi, y) \\ C \end{pmatrix}}_{\Delta} + \underbrace{\begin{pmatrix} (F, e) \\ \Delta \end{pmatrix}}_{\Delta} \begin{pmatrix} \hat{\theta} \\ -1 \end{pmatrix} = 0$$

- $\Delta = (F, e)$ rozszerzona macierz błędu
- Rozmiar macierzy C będzie wynosił $N \times (m + 1)$
- Norma funkcji minimalizującej (*norma Frobeniusa*)

$$\|\Delta\|_F^2 = \sum_{i=1}^M \sum_{j=1}^N \Delta_{ij}^2$$

$$\hat{\theta} = \arg \min \|\Delta\|_F^2$$

Total Least Squares (TLS)

- Macierz C może być wyznaczona jako

$$C = U\Sigma V^T$$

gdzie $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_{n+1})$, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{n+1}$

co jest rozkładem SVD macierzy C

- Macierz \tilde{C} można zapisać jako

$$\tilde{C} = U\Sigma V^T \Leftrightarrow U^{-1}\tilde{C}(V^T)^{-1} = U^T\tilde{C}V = S$$

- Tym samym funkcja celu będzie w postaci

$$\|C - \tilde{C}\|_F^2 = \|U\Sigma V^T - USV^T\|_F^2 = \|\Sigma - S\|_F^2$$

- Macierz S jest diagonalna $S = \text{diag}(s_1, s_2, \dots, s_{n+1})$ i funkcja celu ma postać:

$$V = \|\Sigma - S\|_F^2 = \sum_{i=1}^{m+1} (\sigma_i - s_i)^2$$

$$V = \|\Sigma - S\|_F^2 = \sigma_{n+1}^2$$

Total Least Squares (TLS)

- Ponieważ

$$\mathbf{V} := \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix} \begin{matrix} n \\ d \end{matrix}$$

- Otrzymujemy rozwiązanie

$$\hat{\theta} = -\mathbf{V}_{22}^{-1}\mathbf{V}_{12}$$

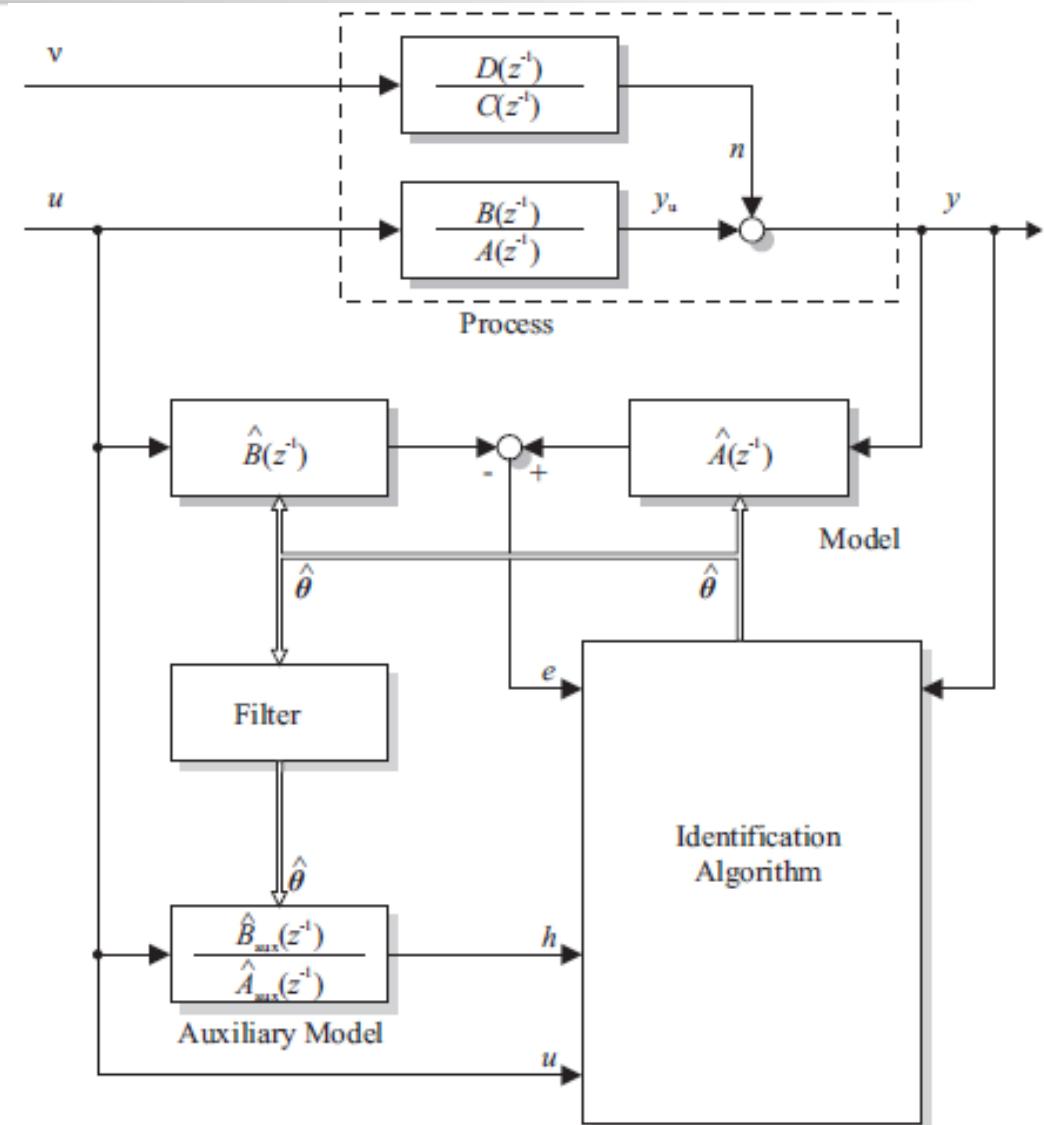
- Metoda TLS dobrze pasuje do identyfikacji obiektów dynamicznych
- TLS jest blisko związana z PCA (Principal Component Analysis) wykorzystywanej w statystyce do wyznaczania korelacji danych i do zmniejszania rozmiaru tych danych.

Metoda Zmiennej Instrumentalnej (IV – Instrumental Variable)

- Metoda bezpośrednia w celu unikania obciążenia estymaty
- Model w postaci $e = y - \Psi\theta$ mnożymy obustronnie przez macierz W – macierz zmiennych instrumentalnych

$$W^T e = W^T y - W^T \Psi\theta$$
- I otrzymujemy rozwiązanie w postaci

$$\hat{\theta} = (W^T \Psi)^{-1} W^T y$$
- Problem jak znaleźć zmienną instrumentalną:
 - Nieskorelowana z $n(k)$
 - Skorelowana z $u(k)$ i $y_u(k)$



Metoda Zmiennej Instrumentalnej (IV – Instrumental Variable)

- Różne metody wyboru zmiennej instrumentalnej:

- Sygnał wejściowy

$$\mathbf{w}^T = (u(k-1-\delta) \dots u(k-m-\delta) | u(k-d-1) \dots u(k-d-m))$$

- Sygnał w oparciu o estymatę niezakłóconego wyjścia $h(k) = \hat{y}_u(k)$

$$\mathbf{w}^T = (-h(k-1) \dots -h(k-m) | u(k-d-1) \dots u(k-d-m))$$

$$h(k) = \hat{y}_u(k) = \boldsymbol{\psi}^T(k) \hat{\boldsymbol{\theta}}(k)$$

- Podejście iteracyjne

1. W pierwszej iteracji stosujemy zmienną instrumentalną w postaci sygnału wejściowego
2. W kolejnej iteracji na podstawie pierwszej iteracji wyniku oraz zmiennej instrumentalnej wygenerowanej w oparciu o estymatę niezakłóconego wyjścia wyliczamy kolejną estymatę parametrów
3. Powtarzamy krok 2 aż parametry już się nie poprawiają

Metody Aproksymacji Stochastycznej (STA – Stochastic Approximation)

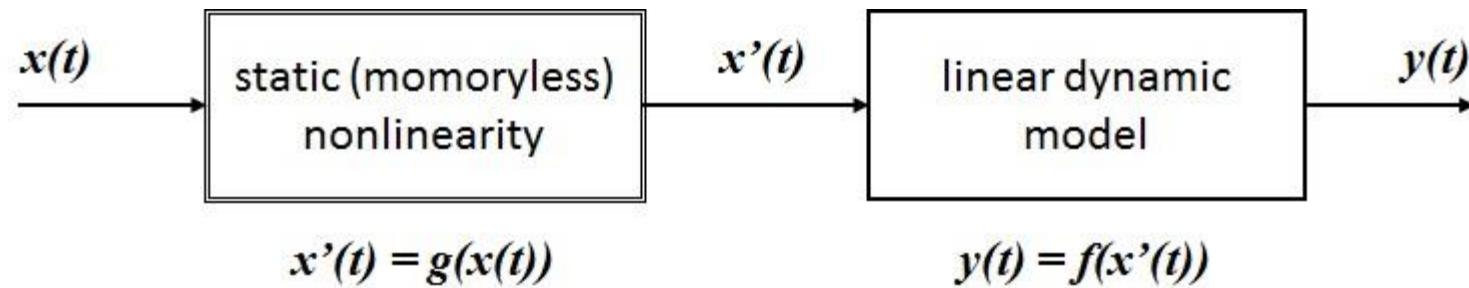
- Metody rekurencyjne
- Mniej wymagające obliczeniowo niż RLS
- Wykorzystujemy gradientowe metody optymalizacji do znalezienia wektora parametrów
 - Robbins-Monro
 - Kiefer-Wolfowitz
- Rzadko stosowane w praktyce, jako że trudno zagwarantować zbieżność metody

Modele nieliniowe

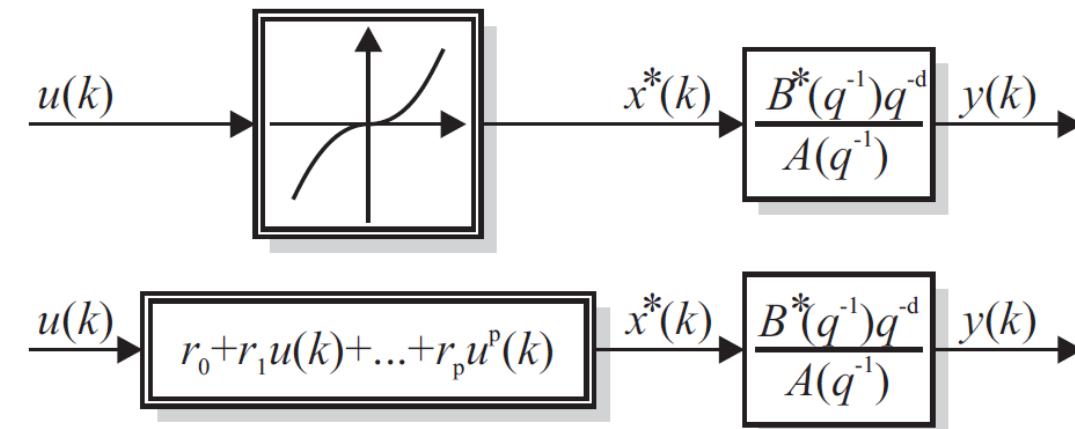
- Model Hammersteina
- Model Wienera
- Szeregi nieliniowe różnych funkcji bazowych:
 - Volter
 - Laguerre
- Modele NARMAX

Model Hammersteina

Wersja ogólna



Wersja szczegółowa - podstawowa



Model Hammersteina podstawowy

- Statyczna nieliniowość w postaci wielomianu

$$x^*(k) = r_0 + r_1 u(k) + r_2 u^2(k) + \dots + r_p u^p(k)$$

- Dynamika liniowa – model ARX

$$A(q^{-1})y(k) = B^*(q^{-1})q^{-d} x^*(k)$$

$$A(q^{-1}) = 1 + a_1 q^{-1} + \dots + a_m q^{-m}$$

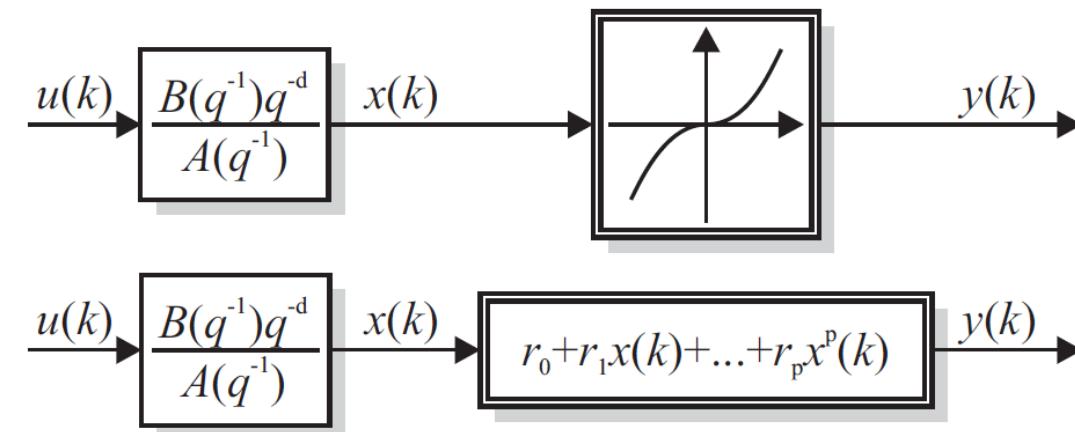
$$B^*(q^{-1}) = b_1^* q^{-1} + \dots + b_m^* q^{-m},$$

Model Hammersteina podstawowy

- Uczenie – proste rozszerzenie metody najmniejszych kwadratów
 - Rozszerzony wektor regresji

$$\psi^T(k) = (-y(k-1), \dots, u(k-1), \dots, u^2(k-1), \dots, u^3(k-1), \dots)$$

Model Wienera



Model Wienera

- Ogólna postać:

$$A_1(q^{-1})y(k) + A_2(q^{-1})y^2(k) + \dots + A_l(q^{-1})y^l(k) = c_{00} + B(q^{-1})u(k-d)$$

- Transmitancja liniowa dynamiczna

$$A(q^{-1})x(k) = B(q^{-1})q^{-d}u(k)$$

- Statyczna nieliniowość wielomianowa

$$y(k) = r_0 + r_1x(k) + r_2x^2(k) + \dots + r_lx^l(k)$$

- Otrzymujemy wersję podstawową

$$y(k) = r_0 + r_1 \frac{B(q^{-1})q^{-d}}{A(q^{-1})} u(k) + r_2 \left(\frac{B(q^{-1})q^{-d}}{A(q^{-1})} \right)^2 u^2(k) + \dots$$

Modele Hammersteina-Wienera

- Optymalizacja nieliniowa
- Uczenie iteracyjne



MI

Metody Identyfikacji

wykład #8

1. *Modele NARMAX*
2. *Szeregi nieliniowe różnych funkcji bazowych:*
 - a) *Volterra*
 - b) *Laguerre*
 - c) ...

Modele NARMAX

- Nieliniowe modele opisane za pomocą równania różnicowego
 - postać ogólna

$$y(k+1) = f(y(k), y(k-1), \dots, u(k), u(k-1), \dots)$$

- Zależność od opóźnionych $y(k)$ – *autoregresja*
- Zależność od opóźnionych $u(k)$ – *ruchoma średnia*
- Wiele wersji, jak na przykład NAR, NMA, NARMA, NARX, etc ...
- Postacie iloczynowe

$$y(k+1) = a_1 y(k) y(k-1) + a_2 y(k) u(k-1) + a_3 u(k) u(k)$$

Modele kawałkami liniowe



Modele Takagi-Sugeno

- W 1985 roku Takagi i Sugeno zaproponowali nowy schemat wnioskowania. Metoda ta zmieniła standardową postać reguł i stanowi rozwinięcie najprostszego układu rozmytego z singletonami po stronie następnika reguły.
- Zakładając, że układ ma N sygnałów wejściowych x_1, x_2, \dots, x_N można zapisać poszczególną regułę w postaci:

R²: **jeśli** x_1 **jest** A1 **i** x_2 **jest** A2 **i** ... x_N **jest** AN **to** $y=f(x_1, x_2, \dots, x_N)$

- gdzie funkcja $f(x_1, x_2, \dots, x_N)$ jest liniową kombinacją wejść

$$f(x_1, x_2, \dots, x_N) = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_N x_N$$

Model NARMAX a model Takagi-Sugeno

- Przyjmujemy opis za pomocą modelu ARX

$$A(z^{-1})y_k = B(z^{-1})z^{-d}u_k + e_k$$

$$A(z^{-1}) = 1 + a_1 z^{-1} + \dots + a_{N_a} z^{-N_a}$$

$$B(z^{-1}) = b_0 + b_1 z^{-1} + \dots + b_{N_b} z^{-N_b}$$

$$Ee_k = 0 \quad Ee_k^2 = \sigma^2$$

- Jedna reguła opisuje model lokalny

$$\hat{y}(t+1) = a_{1,i} y(t) + \dots + a_{N_a,i} y(t - N_a) + b_{0,i} u(t - d + 1) + \dots + b_{N_b,i} u(t - d - N_b)$$

- Natomiast złożenie wszystkich reguł

$$w_i(z(t)) = \frac{|z(t) \text{ jest } Z_i|}{\sum_{j=1}^M |z(t) \text{ jest } Z_j|}$$

- gdzie $|z(t) \text{ jest } Z_i|$ oznacza stopień przynależności $z(t)$ do Z_i .

Model NARMAX a model Takagi-Sugeno

- Model możemy przepisać jako

$$\hat{y}(t+1) = \sum_{i=1}^M \left[a_{1,i} y(t) + \dots + a_{N_a,i} y(t - N_a) + b_{0,i} u(t - d + 1) + \dots + b_{N_b,i} u(t - d - N_b) \right] w_i(z(t))$$

- A uwzględniony zależność na wnioskowanie

$$\hat{y}(t+1) = \frac{\sum_{i=1}^M \left[a_{1,i} y(t) + \dots + a_{N_a,i} y(t - N_a) + b_{0,i} u(t - d + 1) + \dots + b_{N_b,i} u(t - d - N_b) \right] \cdot |z(t) \text{ jest } Z_i|}{\sum_{i=1}^M |z(t) \text{ jest } Z_i|}$$

- Całość można przepisać do postaci jednolitej

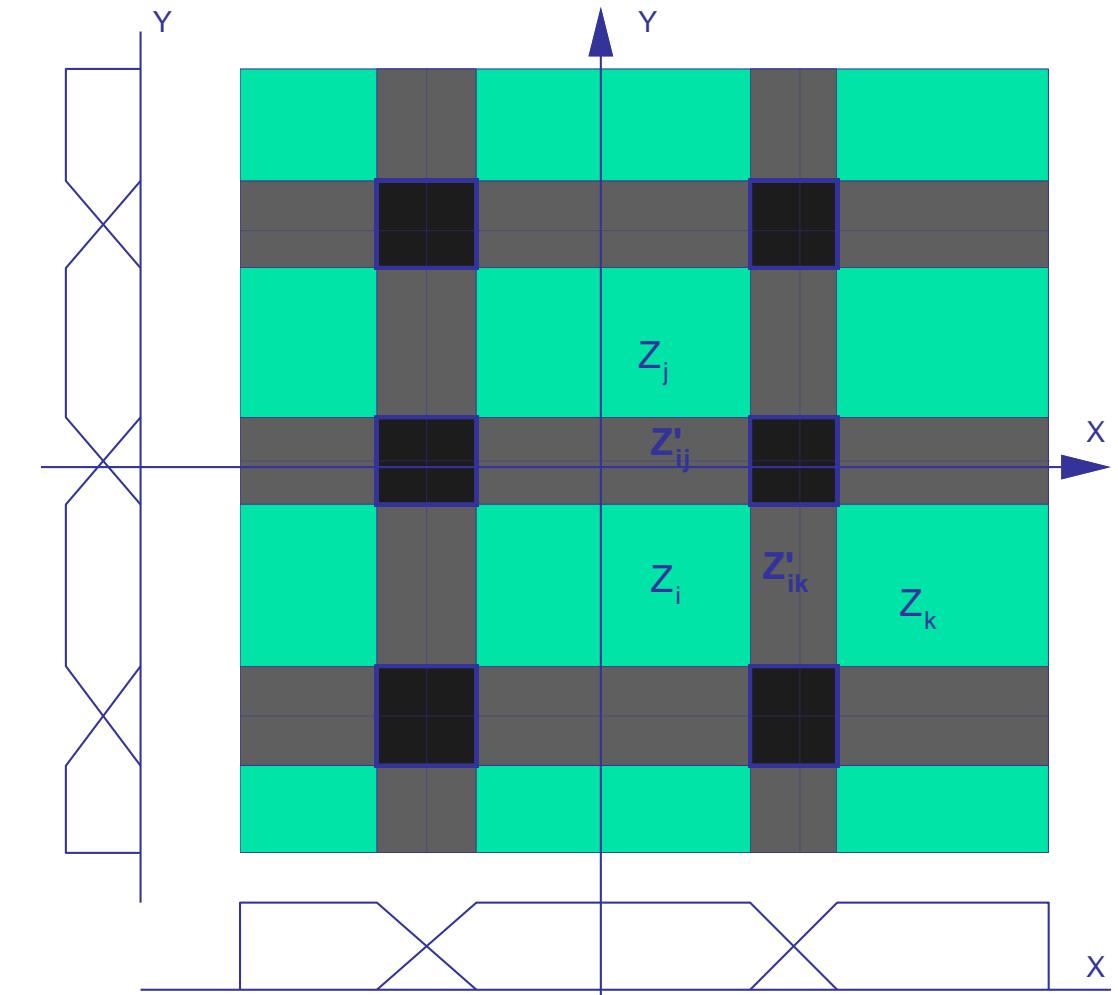
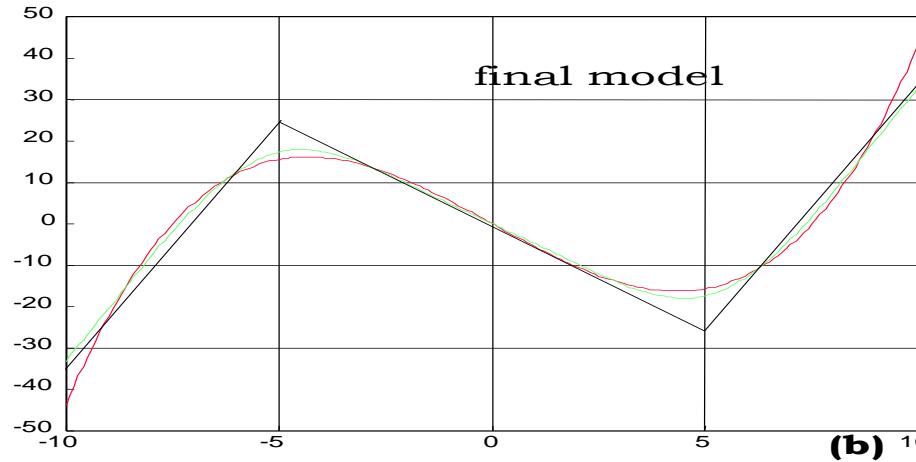
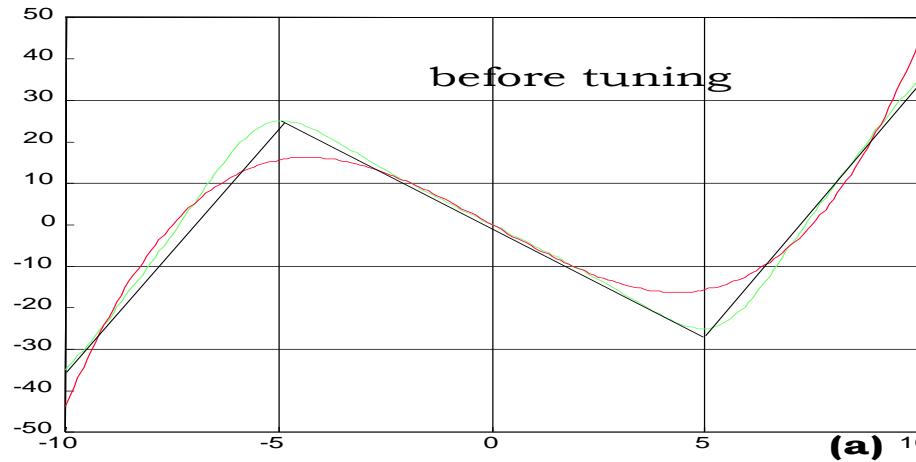
$$y(t+1) = a_1(z(t))y(t) + \dots + a_{N_a}(z(t))y(t - N_a) +$$

$$b_0(z(t))u(t - d + 1) + \dots + b_{N_b}(z(t))u(t - d - N_b)$$

$$a_i(z(t)) = \sum_{j=1}^M a_{i,j} w_j(z(t))$$

$$b_i(z(t)) = \sum_{j=1}^M b_{i,j} w_j(z(t))$$

Rozmyte modelowanie kawałkami liniowe



Modele NARMAX a sieci neuronowe

- Wyobraźmy sobie sieć rekurencyjną
- $y(n+q) = \Phi(\mathbf{x}(n), \mathbf{u}_q(n))$, gdzie q jest rozmiarem przestrzeni stanów i $\Phi: R^{2q} \rightarrow R$.
- Zakładamy, że sieć jest obserwowały, tzn. $\mathbf{x}(n) = \Psi(y_q(n), \mathbf{u}_{q-1}(n))$ where $\Psi: R^{2q} \rightarrow R$.
- $y(n+q) = F(y_q(n), \mathbf{u}_q(n))$ gdzie $\mathbf{u}_{q-1}(n)$ jest zawarte w $\mathbf{u}_q(n)$ jako pierwsze ($q-1$) elementy, a przekształcenie nieliniowe $F: R^{2q} \rightarrow R$ uwzględnia zarówno Ψ jak i Φ .
- $y(n+1) = F(y(n), \dots, y(n-q+1), u(n), \dots, u(n-q+1))$

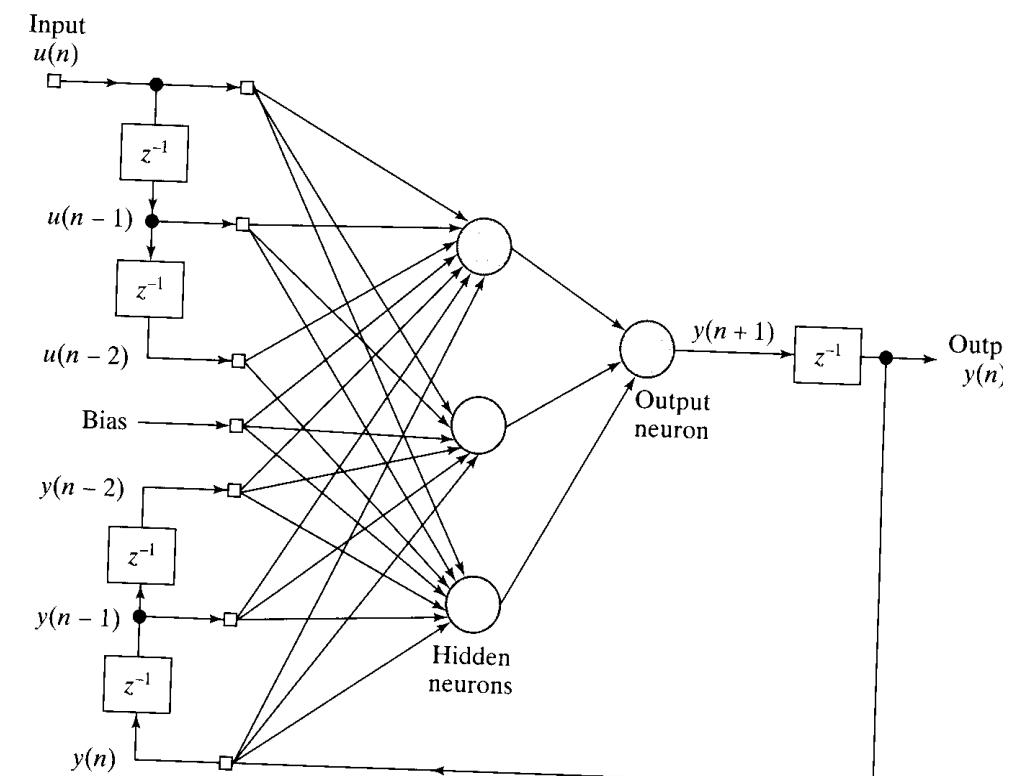
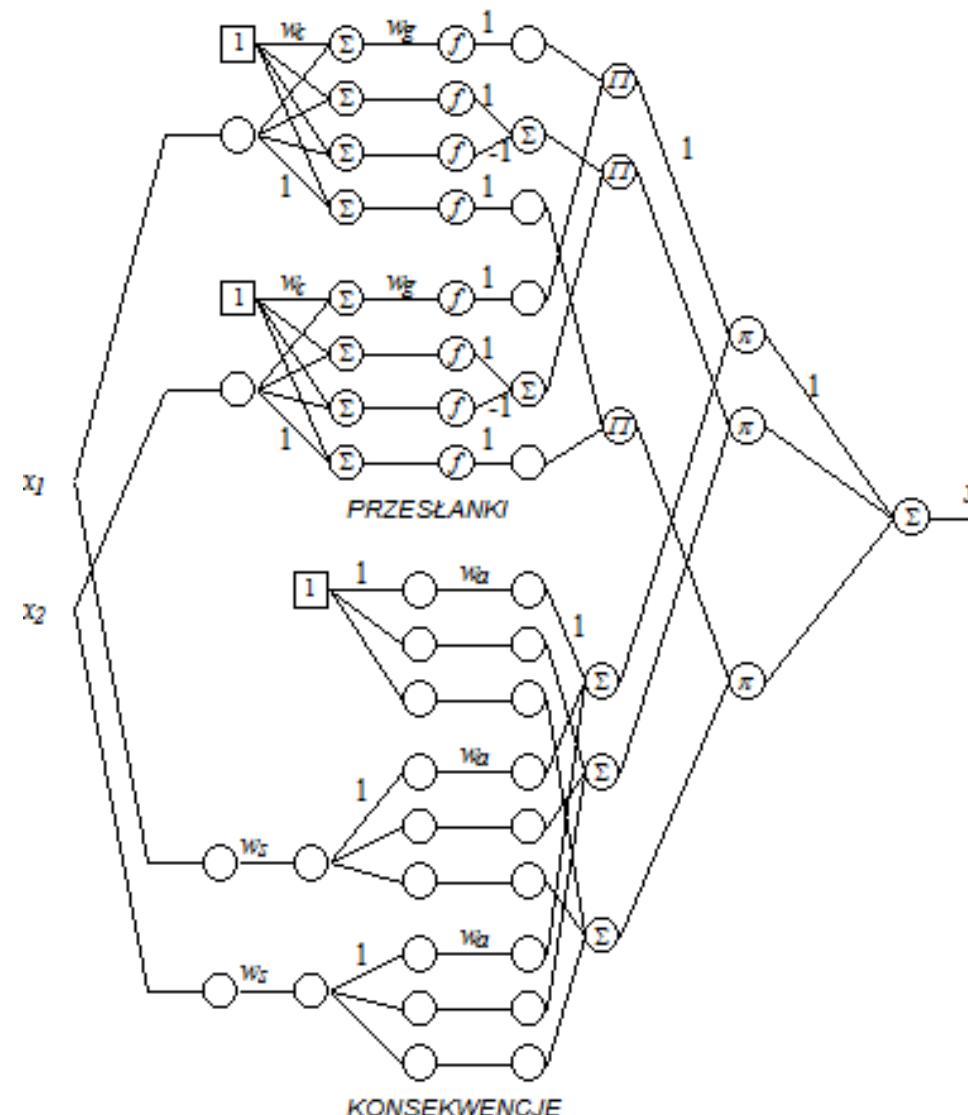


FIGURE 15.7 NARX network with $q = 3$ hidden neurons.

Rozmyte sieci neuronowe

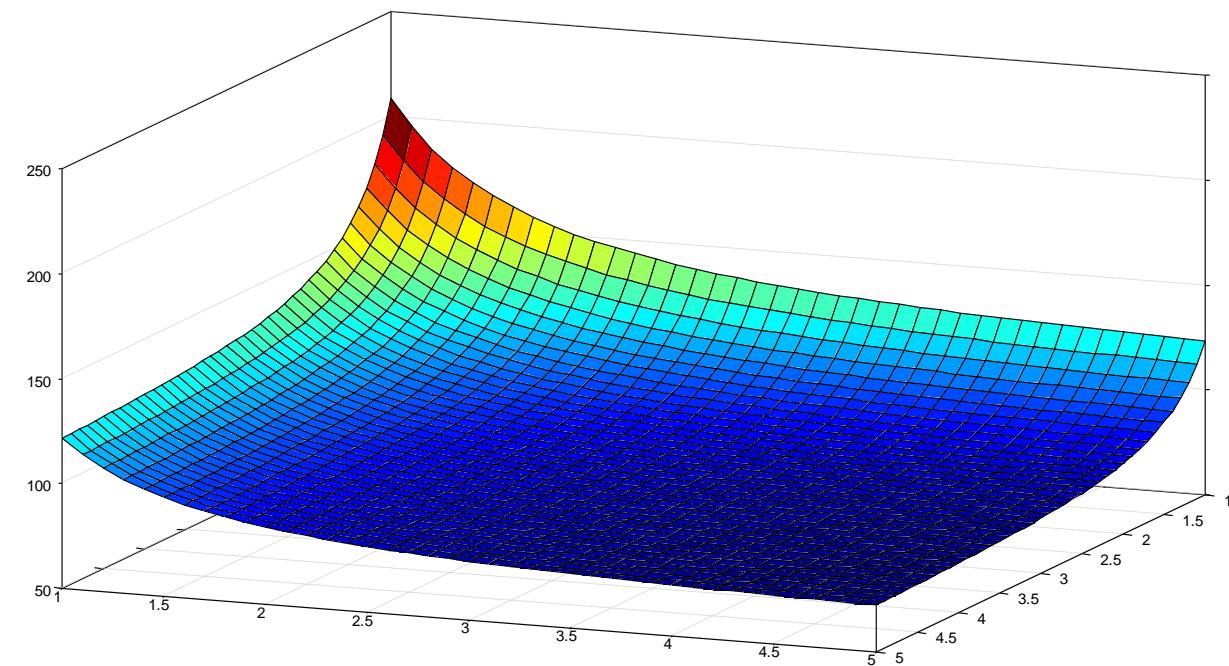


Model NARMAX a model Takagi-Sugeno

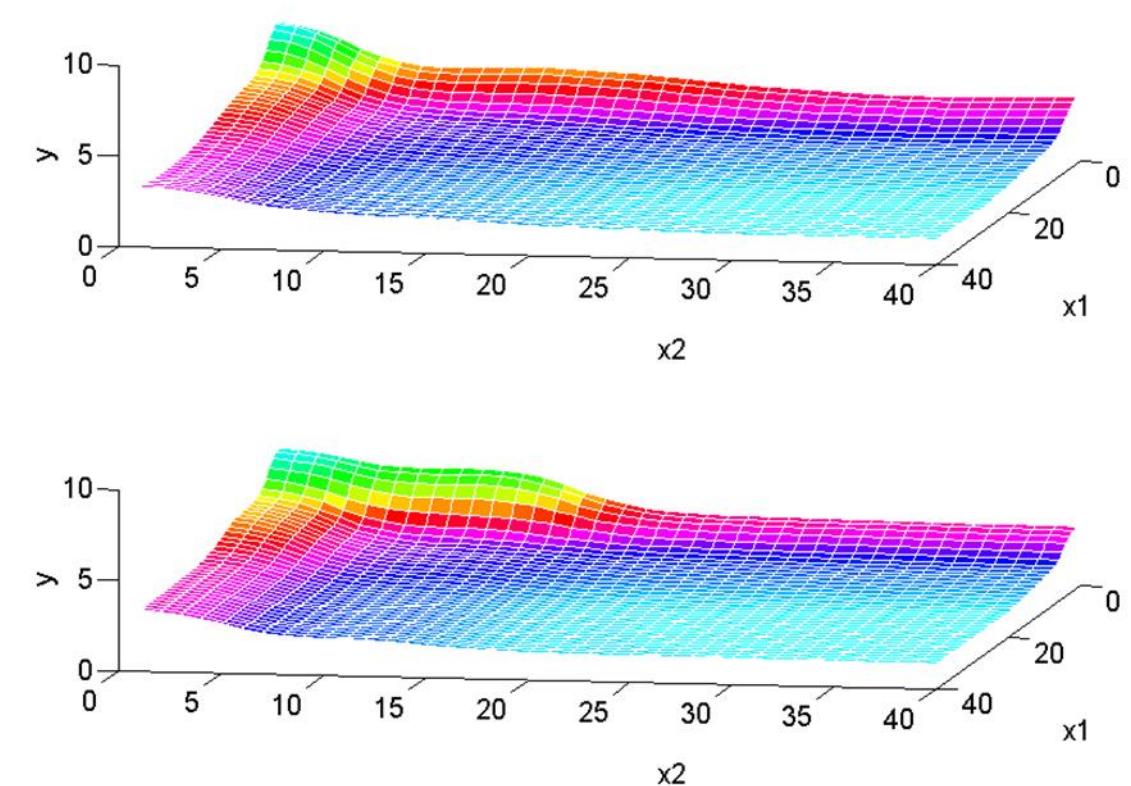
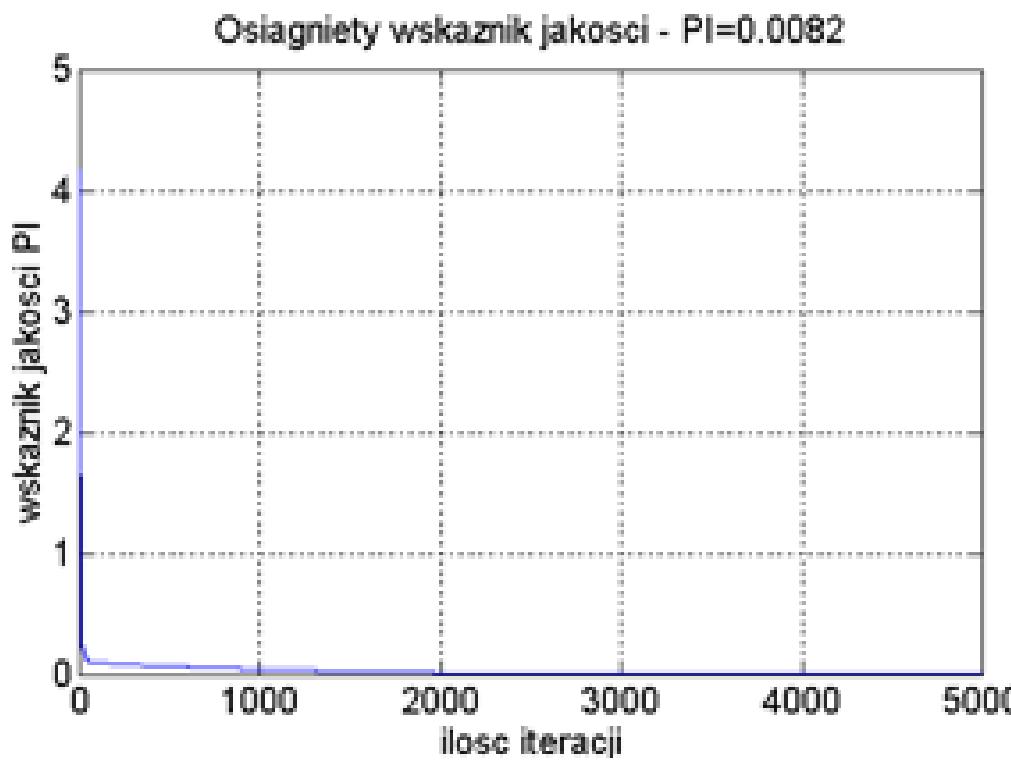
- Przykład: funkcja nieliniowa

$$y = (1 + x_1^{-2} + x_2^{-1.5})^2$$

$$x_1 \in \langle 1,5 \rangle \quad x_2 \in \langle 1,5 \rangle$$



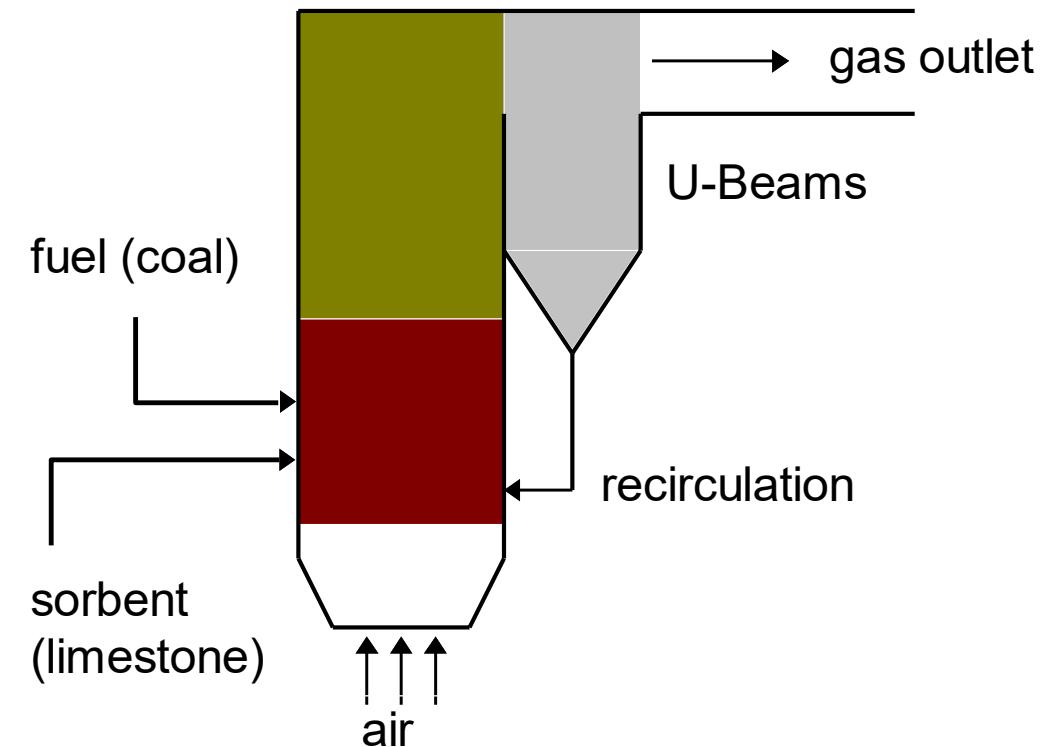
Model NARMAX a model Takagi-Sugeno



Przykład – kocioł fluidalny OFz-450

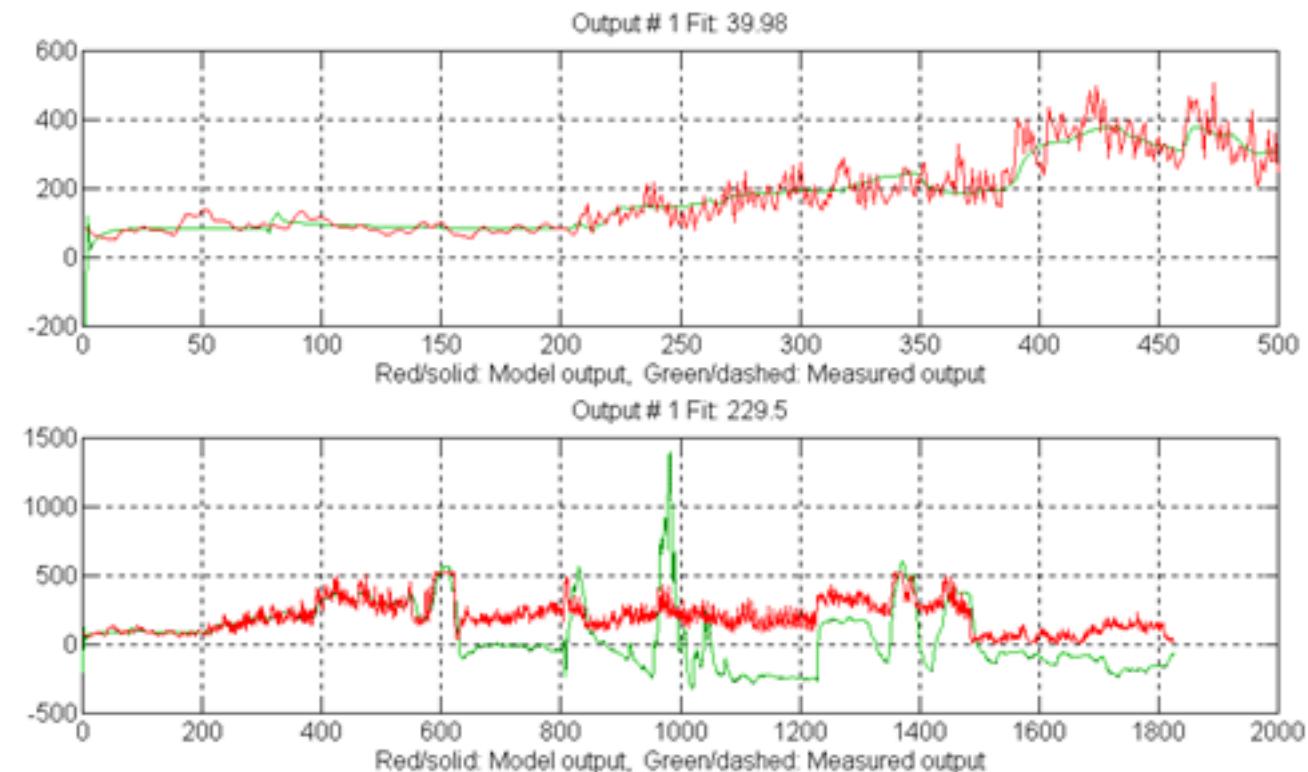
- Przekroczenia emisji SO₂
- Układ sterowania dozowania sorbentu (kamień wapienny)
- Różne modele
 - Model liniowy Box-Jenkins
 - AFNN
 - model jakościowy otrzymany koewolucyjnie
 - model Takagi-Sugeno:
fuzzy neural network (FNN);

PROCESS SCHEMATIC REPRESENTATION



Przykład – kocioł fluidalny OFz-450

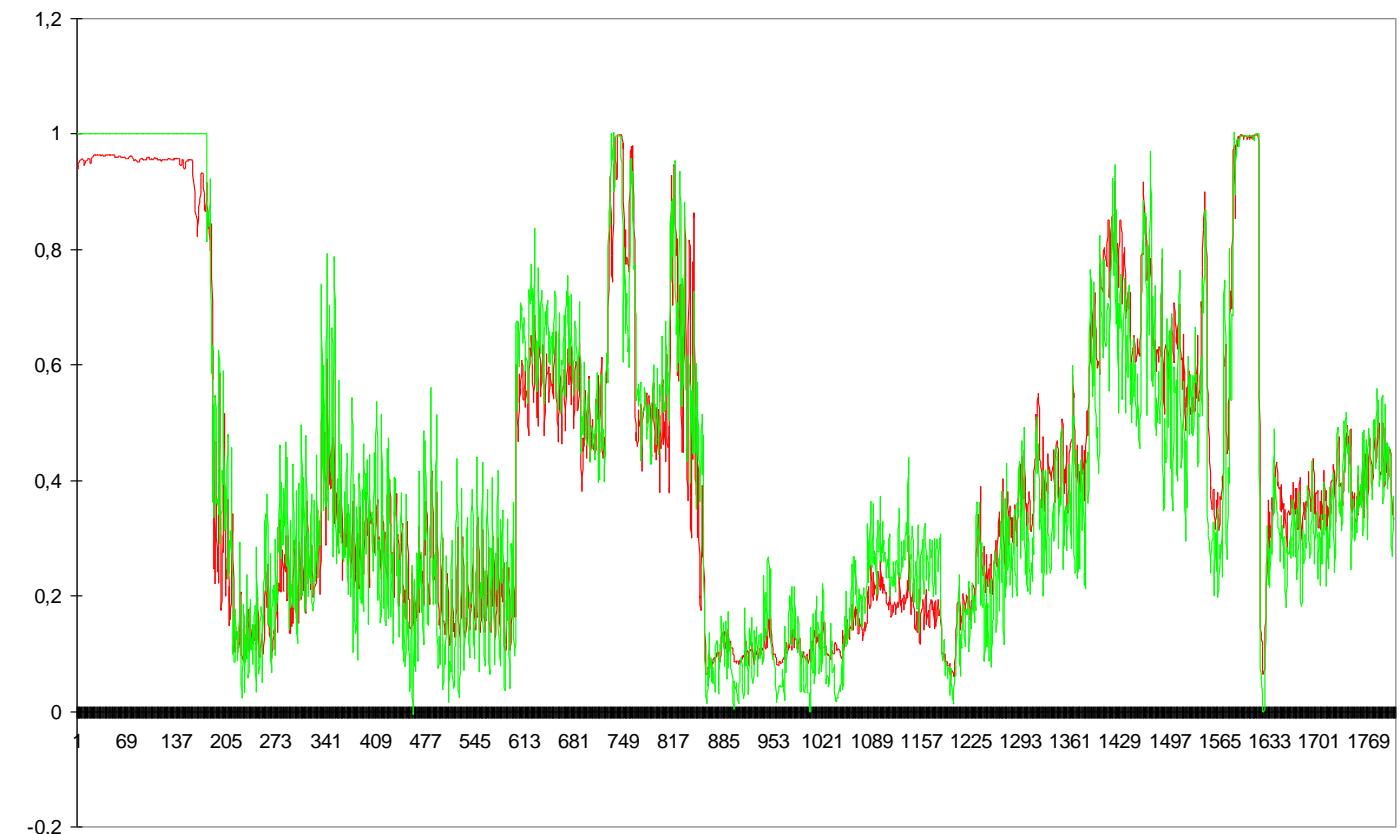
- Model liniowy
- 5 inputs:
steam production;
primary air flow;
sorbent quantity;
temperature at level 2;
temperature at level 4.
- polynomial orders:
 $\{N_A, N_B, N_C, N_D, N_F, d\}$
 $\{1, [2, 2, 5, 1, 1], 0, 0, [3, 5, 6, 2, 2],$
 $[1, 1, 0, 0, 0]\}$



Przykład – kocioł fluidalny OFz-450

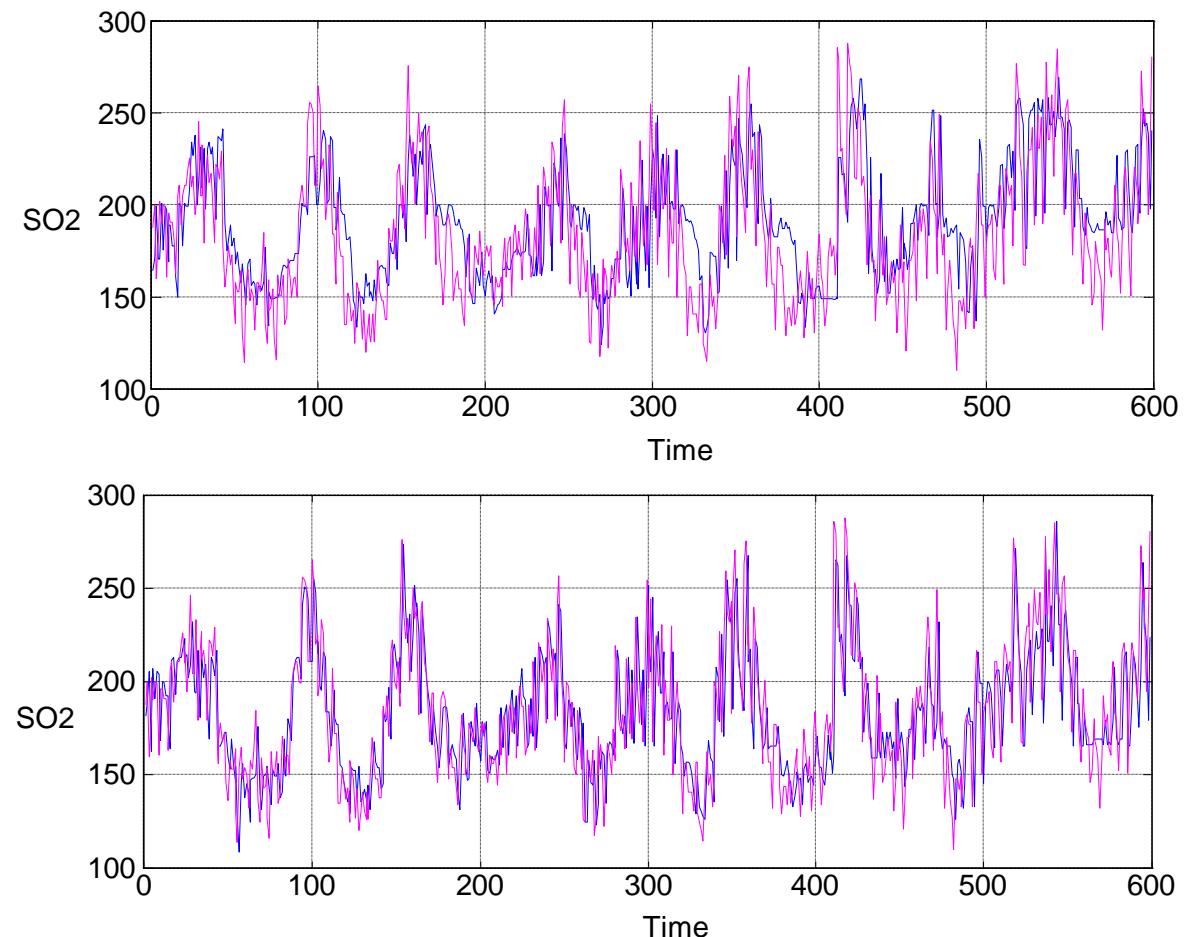
- ANN

- input coal flow;
- primary air flow;
- sorbent quantity;
- temperature at level 1;
- temperature at level 2;
- temperature at level 4.
- past SO_2 emission;
- past CO emission;
- past O_2 concentration.



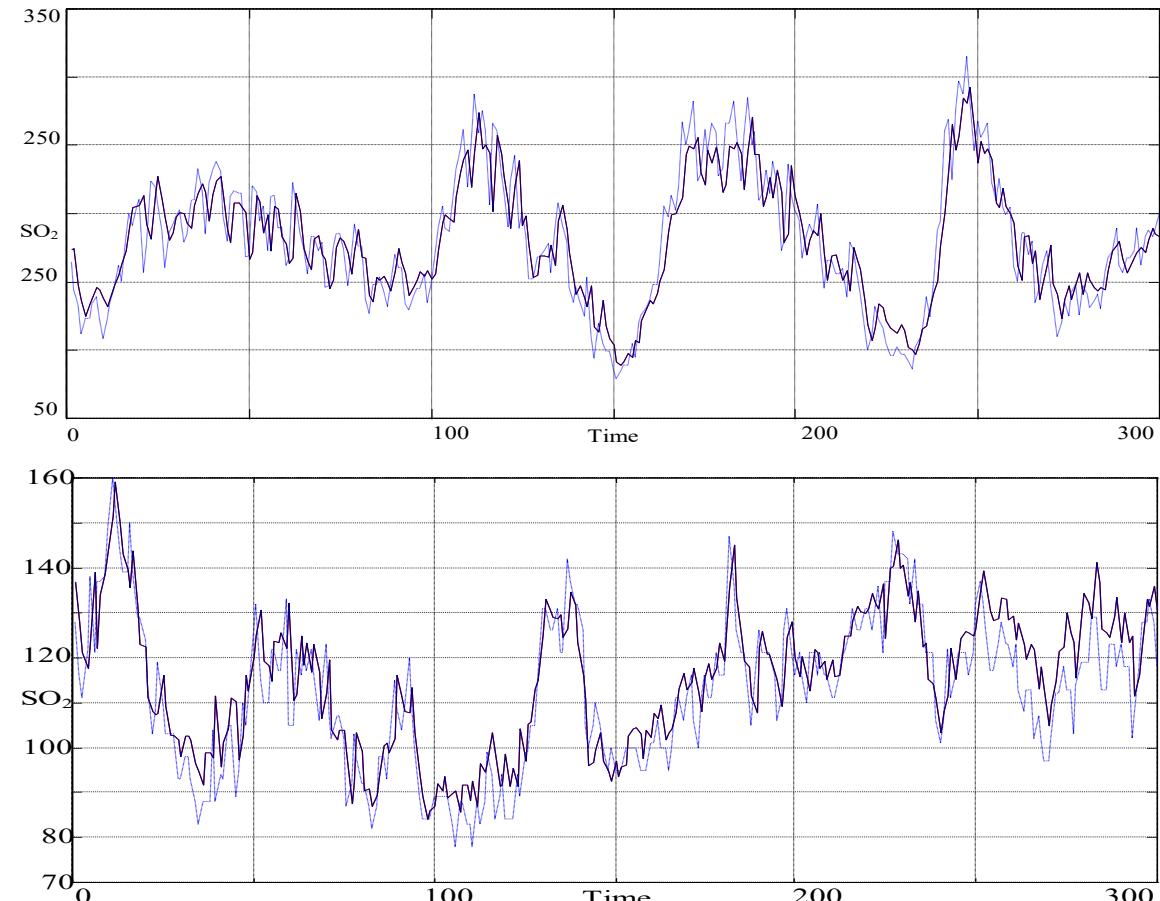
Przykład – kocioł fluidalny OFz-450

- Model koewolucyjny
 - primary air flow;
 - sorbent quantity;
 - boiler load;
 - temperature at level 2;
 - temperature at level 4.
 - past SO_2 emission;



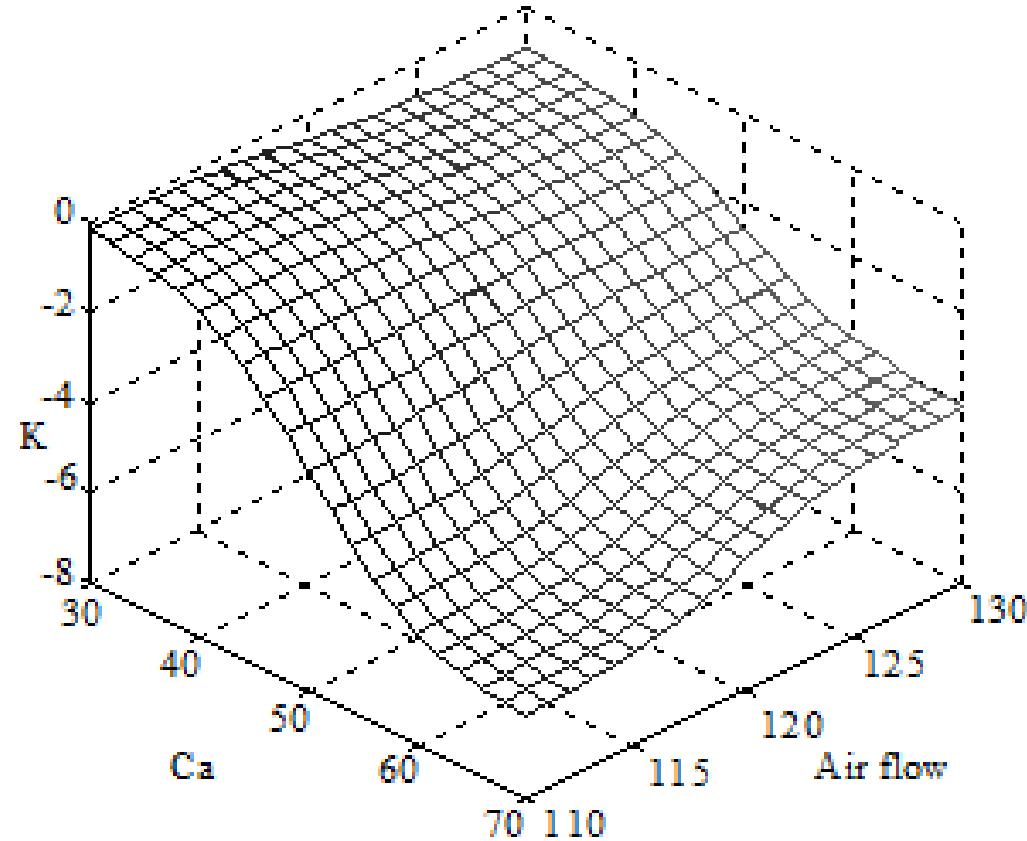
Przykład – kocioł fluidalny OFz-450

- NARMAX: Takagi-Sugeno
 - steam flow;
 - primary air flow;
 - sorbent quantity (k)moment;
 - sorbent quantity ($k-1$)moment;
 - temperature at level 2;
 - temperature at level 4.
 - past SO_2 emission ($k-1$)moment;
 - past SO_2 emission ($k-2$)moment;

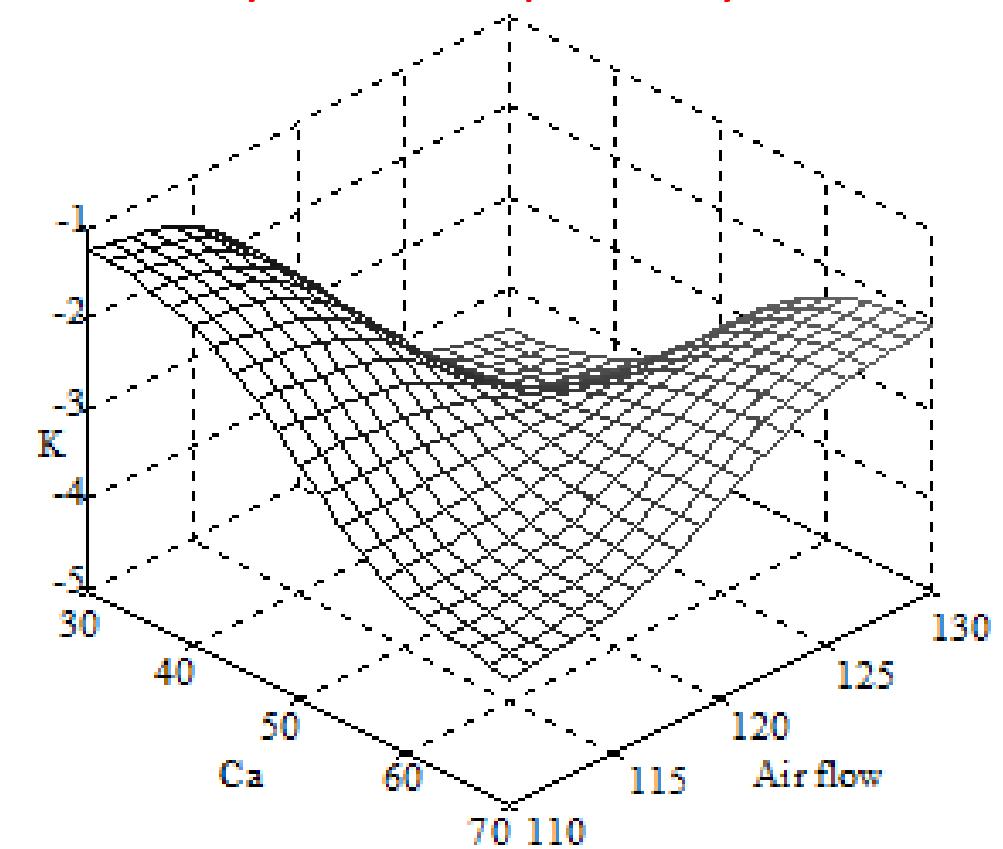


Przykład – kocioł fluidalny OFz-450

niskie temperatury



wysokie temperatury



wzmocnienie $K=SO_2/Ca$ w zależności od dopływu kamienia i powietrza



Szeregi Volterra

- Modele całkowe wysokich rzędów w postaci ciągu całek

- w wersji ciągłej

$$y(t) = \int_{-\infty}^{\infty} h_1(\tau) u(t - \tau) d\tau$$

$$+ \int \int h_2(\tau_1, \tau_2) u(t - \tau_1) u(t - \tau_2) d\tau_1 d\tau_2$$

$$+ \int \int \int h_3(\tau_1, \tau_2, \tau_3) u(t - \tau_1) u(t - \tau_2) u(t - \tau_3) d\tau_1 d\tau_2 d\tau_3$$

+ ...

- w wersji dyskretnej

$$y(k) = \sum_{j=1}^{\infty} h_1(j) u(k - j)$$

$$+ \sum \sum h_2(j_1, j_2) u(k - j_1) u(k - j_2)$$

$$+ \sum \sum \sum h_3(j_1, j_2, j_3) u(k - j_1) u(k - j_2) u(k - j_3)$$

+ ...



Szeregi Volterra

- Modele zawierają jądro (kernel, funkcja bazowa) w postaci $h_3(\tau_1, \tau_2, \dots, \tau_n)$
 - odpowiadające odpowiedzi impulsowej
- Jeśli przyjmiemy, że rząd jest ograniczony do $k < M$, to otrzymamy postać

$$y(k) = c_{00} + \sum_{n=1}^p v_M^n(k)$$

$$v_M^n = \sum_{i_1=0}^M \dots \sum_{i_n=1}^M \alpha_n(i_1, \dots, i_n) u(k - i_1) \dots u(k - i_n)$$

- Inna postać ograniczona do rzędu p

$$A(q^{-1})y(k) = c_{00} + B_1(q^{-1})u(k - d)$$

$$+ \sum_{\beta_1=0}^h B_{2\beta_1}(q^{-1})u(k - d)u(k - d - \beta_1) + \dots$$

$$+ \sum_{\beta_1=0}^h \sum_{\beta_2=\beta_1}^h \dots \sum_{\beta_{p-1}=\beta_{p-2}}^h B_{p\beta_1\beta_2\dots\beta_{p-1}}(q^{-1})u(k - d) \prod_{\xi=1}^{p-1} u(k - d - \beta_\xi) + \dots$$

często ograniczane do rzędu drugiego

Aproksymacja za pomocą wielomianów ortogonalnych

- Postać wielomianu aproksymującego

$$Q(x) = \sum_{k=0}^m c_k \varphi_k(x)$$

- Ciąg funkcji jest ortogonalny gdy:

$$\sum_{i=0}^n \varphi_k(x_i) \cdot \varphi_l(x_i) = 0 \text{ dla } k, l = 0, 1, \dots, m, k \neq l$$

Wielomiany Legendre'a

- Postać rekurencyjna

$$L_0(x) = 1$$

$$L_1(x) = x$$

$$L_{n+1}(x) = \left(\frac{2n+1}{n+1} \right) x L_n(x) - \left(\frac{n}{n+1} \right) L_{n-1}(x) \quad n=1,2,\dots$$

- Wielomian Legendre'a

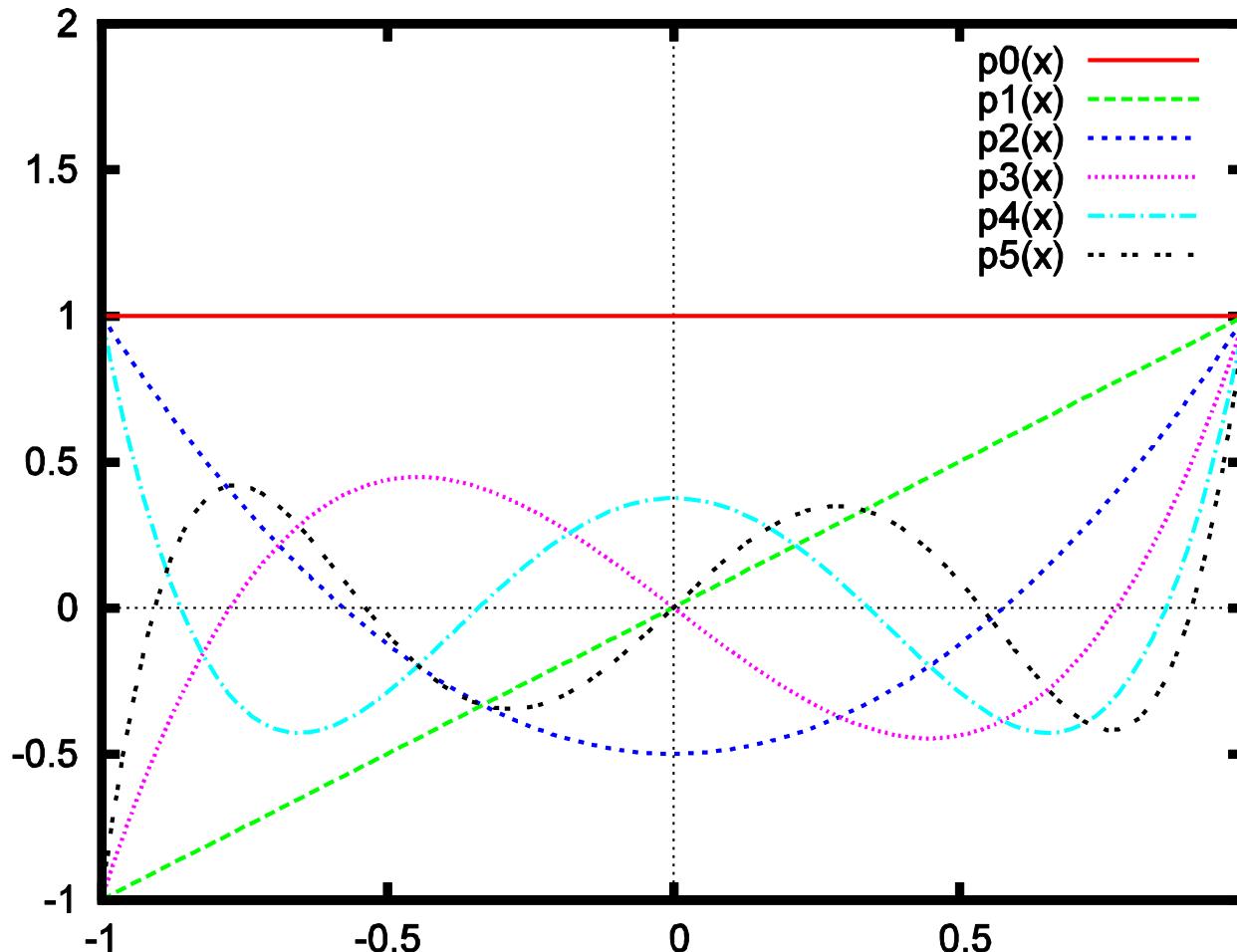
$$P_k(x) = \frac{1}{2^k \cdot k!} \cdot \frac{d^k}{dx^k} (x^2 - 1)^k$$

- Wielomiany są ortogonalne w dziedzinie $x \in [-1,1]$

$$C_k = \frac{\sum_{i=0}^n y_i \cdot P_k(x_i)}{\sum_{i=0}^n P_k^2(x_i)}$$



5 pierwszych wielomianów Legendre'a



$$P_0(x) = 1$$

$$P_1(x) = x$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1)$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x)$$

$$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$$

$$P_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x)$$

Wielomiany Czebyszewa

Wzór ogólny

$$T_0(x) = 1$$

$$T_1(x) = x$$

$$T_2(x) = 2x^2 - 1$$

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x)$$

$$T_0(x) = 1$$

$$T_1(x) = x$$

$$T_2(x) = 2x^2 - 1$$

$$T_3(x) = 4x^3 - 3x$$

$$T_4(x) = 8x^4 - 8x^2 + 1$$

$$T_5(x) = 16x^5 - 20x^3 + 5x$$

$$T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1$$

$$T_7(x) = 64x^7 - 112x^5 + 56x^3 - 7x$$



$$U_0(x) = 1$$

$$U_1(x) = 2x$$

$$U_2(x) = 4x^2 - 1$$

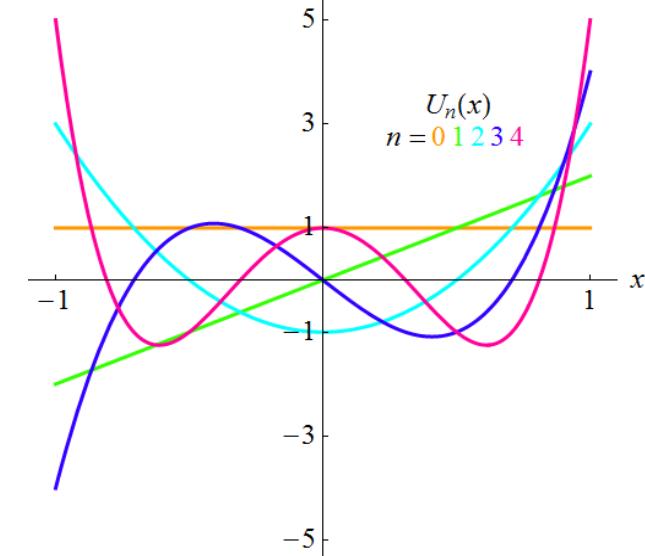
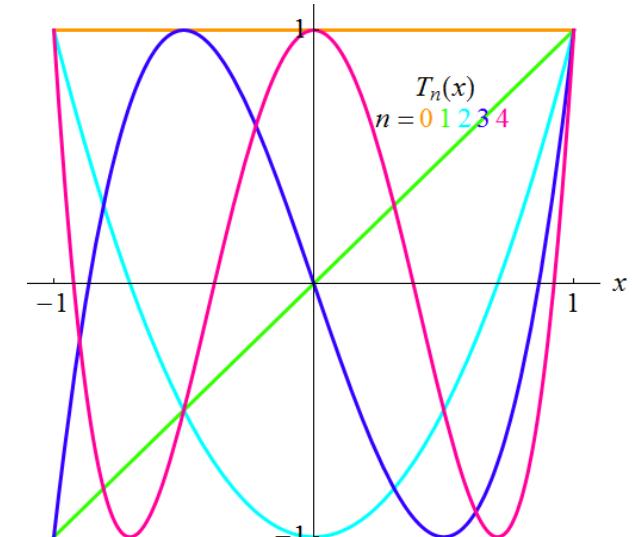
$$U_3(x) = 8x^3 - 4x$$

$$U_4(x) = 16x^4 - 12x^2 + 1$$

$$U_5(x) = 32x^5 - 32x^3 + 6x$$

$$U_6(x) = 64x^6 - 80x^4 + 24x^2 - 1$$

$$U_7(x) = 128x^7 - 192x^5 + 80x^3 - 8x$$



Wielomiany Laguerre'a

$$L_n(x) = \frac{e^x}{n!} \frac{d^n}{dx^n} (x^n e^{-x}) \rightarrow L_n \text{ ortogonalny na } [0, \infty]$$

$$(n+1)L_{n+1}(x) = (2n+1-x)L_n(x) - nL_{n-1}(x)$$

$$L_0^{(\alpha)}(x) = 1$$

$$L_1^{(\alpha)}(x) = -x + \alpha + 1$$

$$L_2^{(\alpha)}(x) = \frac{1}{2}x^2 - (\alpha + 2)x + \frac{(\alpha + 1)(\alpha + 2)}{2}$$

$$L_3^{(\alpha)}(x) = -\frac{1}{6}x^3 + \frac{\alpha + 3}{2}x^2 - \frac{(\alpha + 2)(\alpha + 3)}{2}x + \frac{(\alpha + 1)(\alpha + 2)(\alpha + 3)}{6}$$

$$0! L_0(x) = 1$$

$$1! L_1(x) = 1 - x$$

$$2! L_2(x) = x^2 - 4x + 2$$

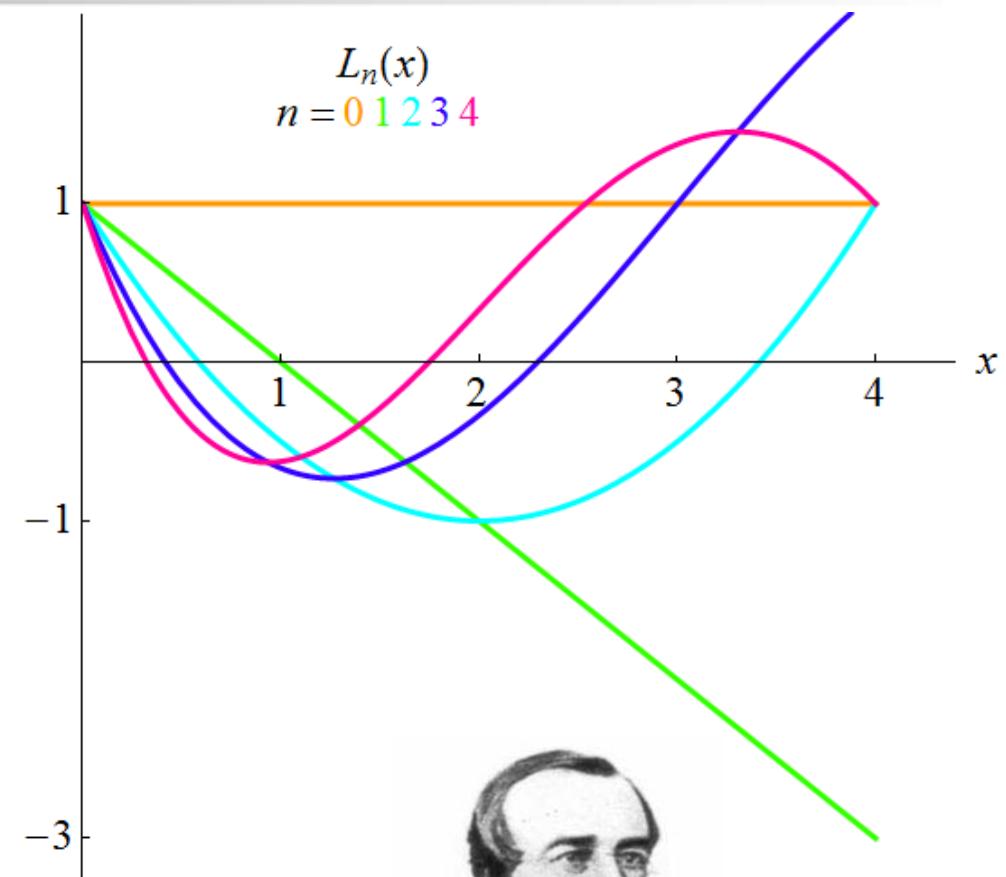
$$3! L_3(x) = -x^3 + 9x^2 - 18x + 6$$

$$4! L_4(x) = x^4 - 16x^3 + 72x^2 - 96x + 24$$

$$5! L_5(x) = -x^5 + 25x^4 - 200x^3 + 600x^2 - 600x + 120$$

$$6! L_6(x) = x^6 - 36x^5 + 450x^4 - 2400x^3 + 5400x^2 - 4320x + 720$$

$$7! L_7(x) = -x^7 + 49x^6 - 882x^5 + 7350x^4 - 29400x^3 + 52920x^2 - 35280x + 5040$$





MI

Metody Identyfikacji

wykład #9

1. *Identyfikacja charakterystyk statycznych*
2. *Regresja liniowa*
3. *Regresja nieliniowa*
4. *Wygładzanie*
5. *Estymatory jądrowe – kernel estimation*
6. *Zadanie klasyfikacji a zadanie interpolacji*

Wstęp

- O co walczymy dokąd zmierzamy? – różne hasła → te same metody
 - Machine Learning
 - Data Mining
 - Pattern Recognition
 - Data Analysis
 - Statistics
- Automatyczna identyfikacja nieprzypadkowych struktur w danych
- Czego oczekujemy:
 - Algorytmu odpornego na wartości poboczne (odległe) i błędne założenia odnośnie modelu.
 - Algorytmu stabilnego: dobra generalizacja w przypadku nowych danych.
 - Algorytmu wydajnego obliczeniowo: wiele danych.

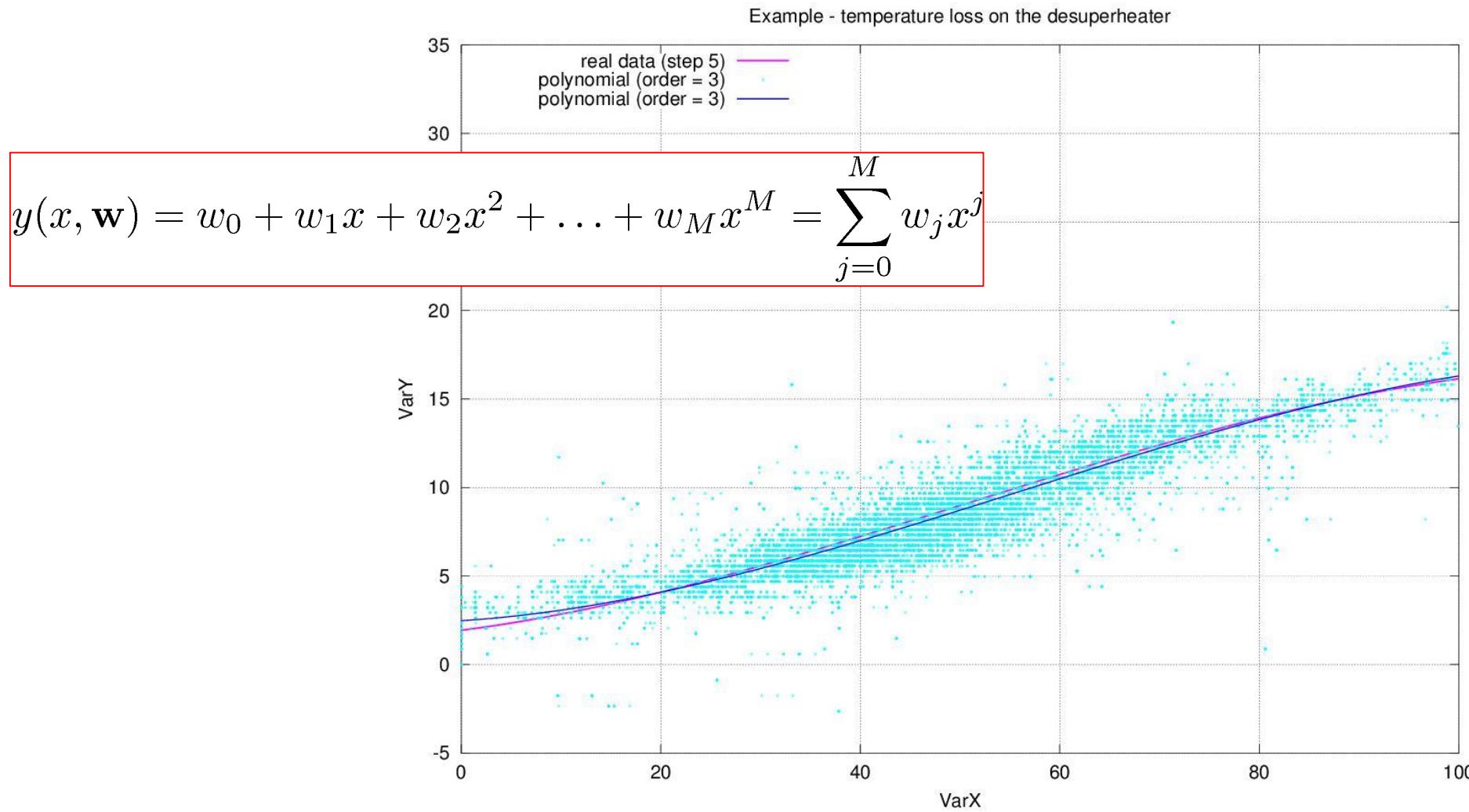
Motywacja

- Wszechstronność w analizie pojawiających się w danych zależności
- Możliwość predykcji zachowania bez konieczności budowania i identyfikacji modelu parametrycznego
- Jak wykryć błędne zachowanie / informację / daną analizując lokalne punkty
- Elastyczność w odtwarzaniu brakujących danych oraz w zadaniu interpolacji

Zakres

- ... zaczynamy od **regresji liniowej**
- Regresja nieliniowa
- Wygładzanie
- Estymacja jądrowa

Regresja liniowa – dopasowanie funkcji wielomianowej

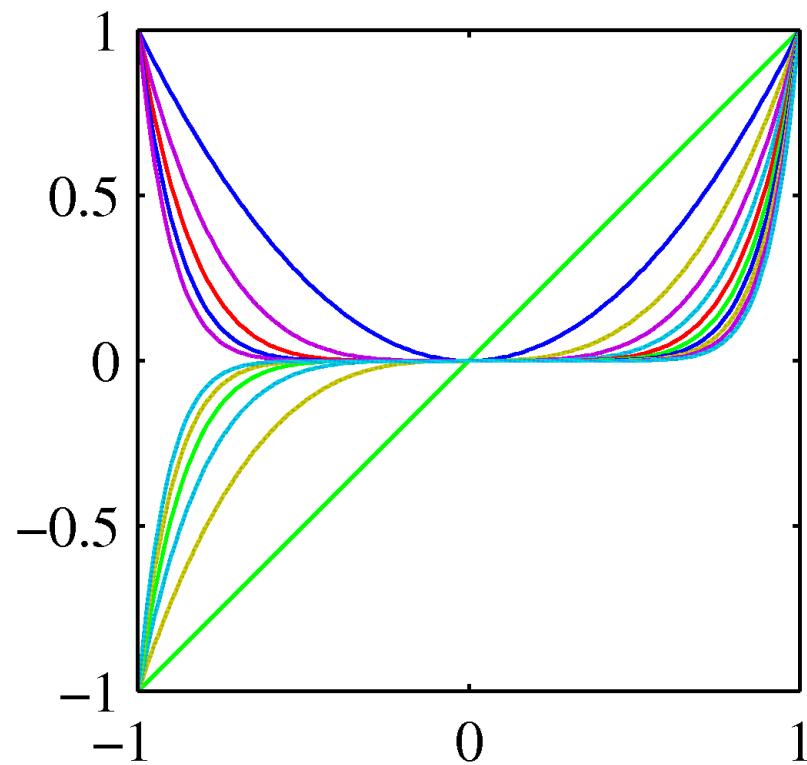


Modele regresji liniowej

- Wielomianowe funkcje bazowe:

$$\phi_j(x) = x^j.$$

- Są globalne: mała zmiana w x zmienia wszystkie funkcje bazowe.

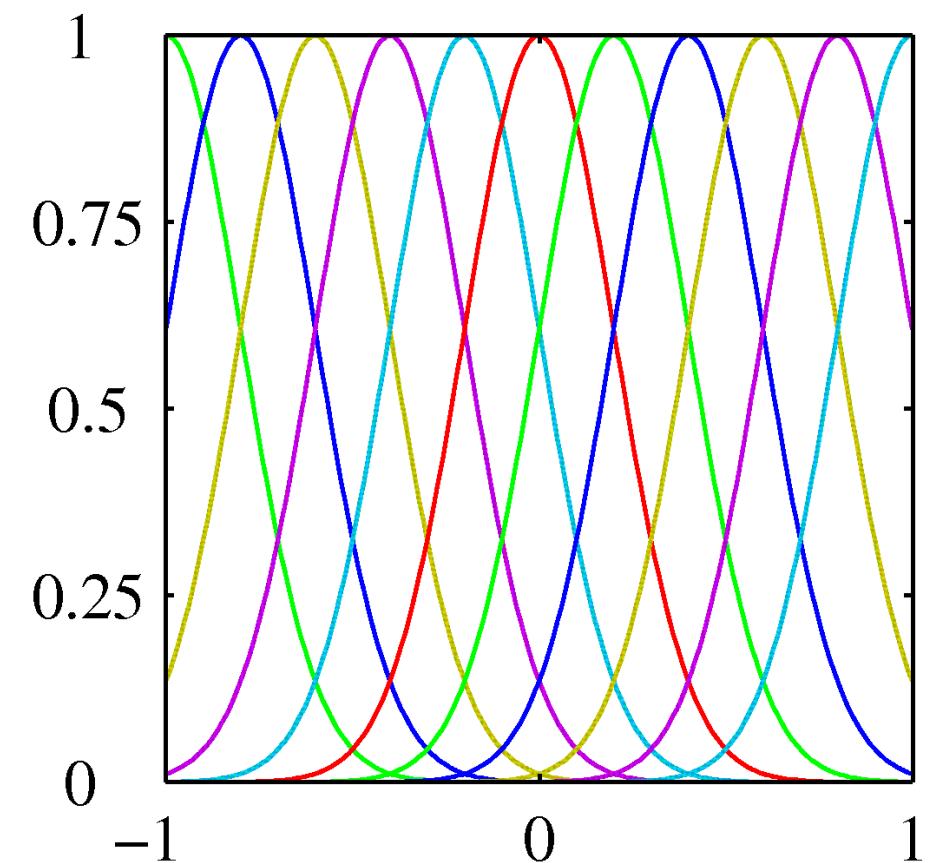


Gaussowskie funkcje bazowe

- Funkcja Gaussa:

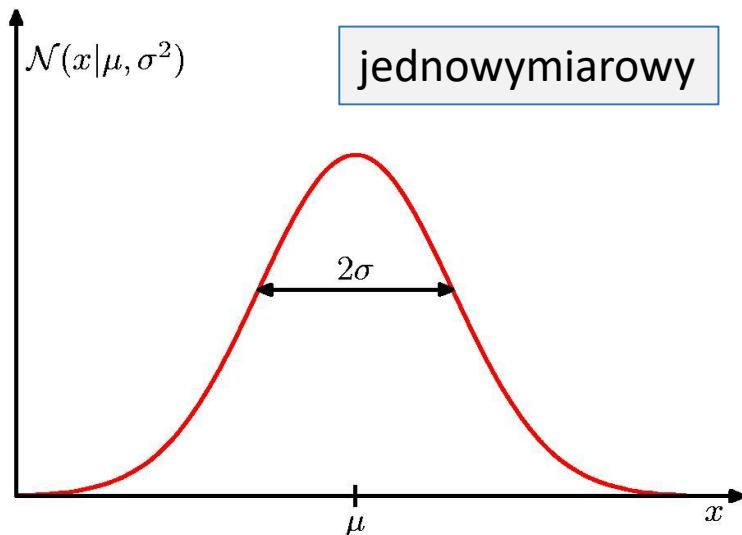
$$\phi_j(x) = \exp \left\{ -\frac{(x - \mu_j)^2}{2s^2} \right\}$$

- Lokalne: mała zmiana w x zmienia tylko okoliczne (lokalne) funkcje.
- Metoda jądrowa

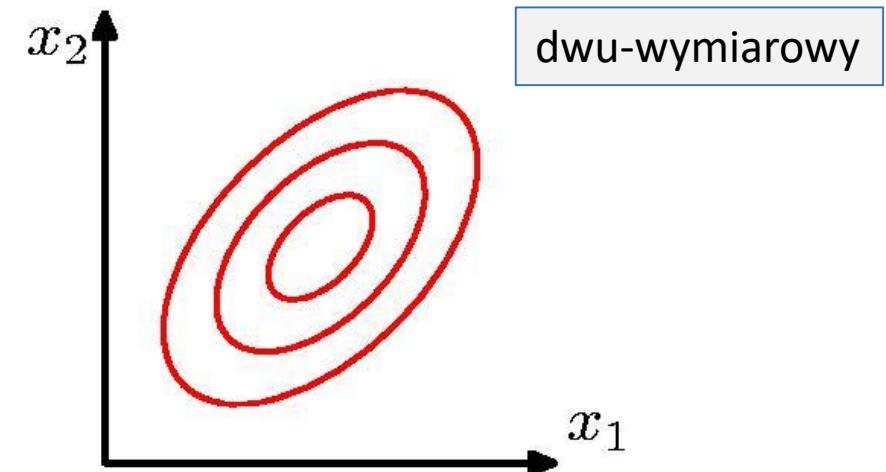


Rozkład Gaussa

$$\mathcal{N}(x|\mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp \left\{ -\frac{1}{2\sigma^2}(x - \mu)^2 \right\}$$



jednowymiarowy



dwu-wymiarowy

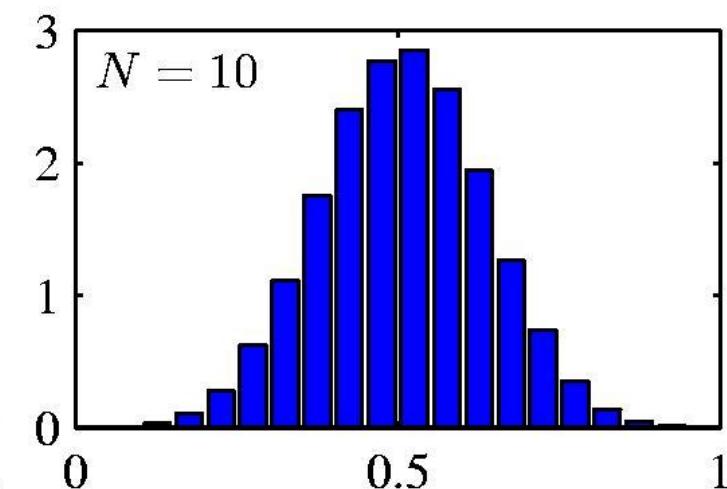
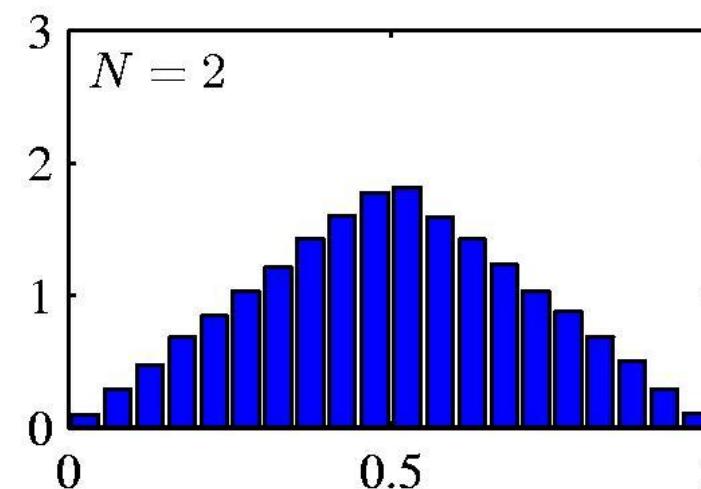
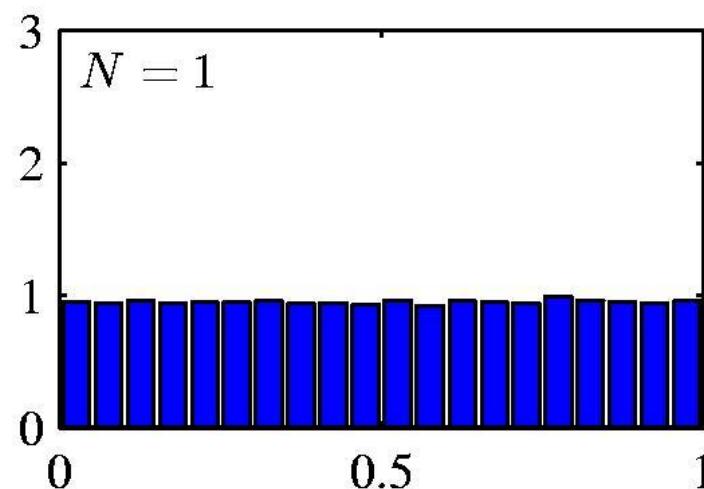
wielowymiarowy



$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp \left\{ -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}$$

Centralne twierdzenie graniczne

- Jeśli zmienne są niezależnymi zmiennymi losowymi o jednakowym rozkładzie, takiej samej wartości oczekiwanej i skończonej wariancji, to zmienna losowa w postaci sumy lokalnych zmiennych losowych zbiega według rozkładu do standardowego rozkładu normalnego, gdy N rośnie do nieskończoności.



Identyfikacja (1)

Maximum Likelihood / Least Squares

- Założymy obserwacje funkcji deterministycznej z dodanym szumem Gaussowskim:

$$t = y(\mathbf{x}, \mathbf{w}) + \epsilon \quad \text{gdzie} \quad p(\epsilon|\beta) = \mathcal{N}(\epsilon|0, \beta^{-1})$$

- Co jest równoważne z ,

$$p(t|\mathbf{x}, \mathbf{w}, \beta) = \mathcal{N}(t|y(\mathbf{x}, \mathbf{w}), \beta^{-1}).$$

- Dla zaobserwowanych wejść $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ i wyjść $\mathbf{t} = [t_1, \dots, t_N]^T$ otrzymujemy ocenę wiarygodności

$$p(\mathbf{t}|\mathbf{X}, \mathbf{w}, \beta) = \prod_{n=1}^N \mathcal{N}(t_n | \mathbf{w}^T \phi(\mathbf{x}_n), \beta^{-1}).$$

Identyfikacja (2)

Maximum Likelihood / Least Squares

- Z algorytmu otrzymujemy

$$\begin{aligned}\ln p(\mathbf{t}|\mathbf{w}, \beta) &= \sum_{n=1}^N \ln \mathcal{N}(t_n | \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_n), \beta^{-1}) \\ &= \frac{N}{2} \ln \beta - \frac{N}{2} \ln(2\pi) - \beta E_D(\mathbf{w})\end{aligned}$$

- gdzie

$$E_D(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{t_n - \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_n)\}^2$$

- jest sumą kwadratów błędów.

Identyfikacja (3)

Maximum Likelihood / Least Squares

- Wyznaczając gradient i przyrównując go do zera otrzymujemy

$$\nabla_{\mathbf{w}} \ln p(\mathbf{t}|\mathbf{w}, \beta) = \beta \sum_{n=1}^N \{t_n - \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_n)\} \boldsymbol{\phi}(\mathbf{x}_n)^T = \mathbf{0}.$$

- A rozwiązuje w względem w otrzymujemy

$$\mathbf{w}_{ML} = \underbrace{\left(\boldsymbol{\Phi}^T \boldsymbol{\Phi} \right)^{-1}}_{\text{Pseudo-odwrotność, Moore-Penrose } \boldsymbol{\Phi}^\dagger} \boldsymbol{\Phi}^T \mathbf{t}$$

- gdzie

$$\boldsymbol{\Phi} = \begin{pmatrix} \phi_0(\mathbf{x}_1) & \phi_1(\mathbf{x}_1) & \cdots & \phi_{M-1}(\mathbf{x}_1) \\ \phi_0(\mathbf{x}_2) & \phi_1(\mathbf{x}_2) & \cdots & \phi_{M-1}(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_0(\mathbf{x}_N) & \phi_1(\mathbf{x}_N) & \cdots & \phi_{M-1}(\mathbf{x}_N) \end{pmatrix}.$$

Identyfikacja (4)

Maximum Likelihood / Least Squares

- Minimalizując tylko względem wyrazu wolnego w_0 widzimy, że:

$$\begin{aligned}
 w_0 &= \bar{t} - \sum_{j=1}^{M-1} w_j \bar{\phi}_j \\
 &= \underbrace{\frac{1}{N} \sum_{n=1}^N t_n}_{\bar{t}} - \sum_{j=1}^{M-1} w_j \underbrace{\frac{1}{N} \sum_{n=1}^N \phi_j(\mathbf{x}_n)}_{\bar{\phi}_j}.
 \end{aligned}$$

- We Maksymalizując względem \mathbf{w} daje:

$$\frac{1}{\beta_{ML}} = \frac{1}{N} \sum_{n=1}^N \{t_n - \mathbf{w}_{ML}^T \boldsymbol{\phi}(\mathbf{x}_n)\}^2$$

Regularyzacja kwadratowa

- Karzemy współczynniki o wysokich wartościach

$$\tilde{E}(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2 + \frac{\lambda}{2} \|\mathbf{w}\|^2$$

Regularyzowana MNK (1)

- Rozważmy funkcje błędu:

$$E_D(\mathbf{w}) + \lambda E_W(\mathbf{w})$$

Dane + Regularyzacja

- Otrzymujemy wskaźnik jakości w postaci

$$\frac{1}{2} \sum_{n=1}^N \{t_n - \mathbf{w}^T \phi(\mathbf{x}_n)\}^2 + \frac{\lambda}{2} \mathbf{w}^T \mathbf{w}$$

- I rozwiązanie

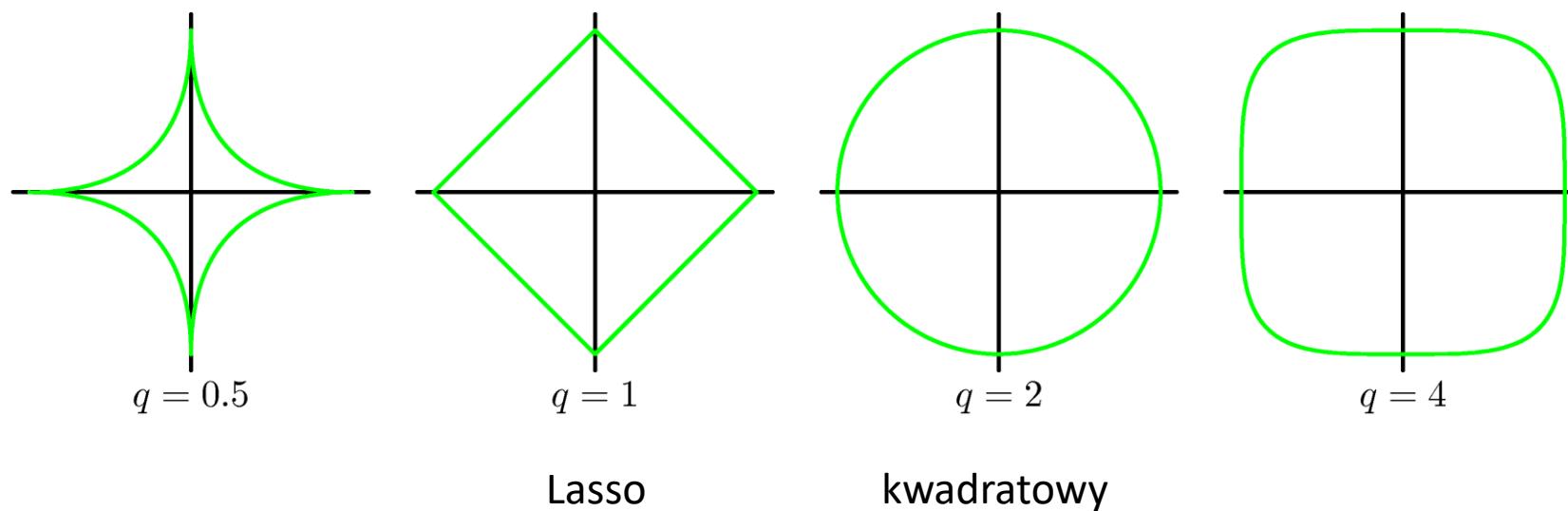
$$\mathbf{w} = (\lambda \mathbf{I} + \Phi^T \Phi)^{-1} \Phi^T \mathbf{t}.$$

λ : współczynnik
regularizacji

Regularyzowana MNK (2)

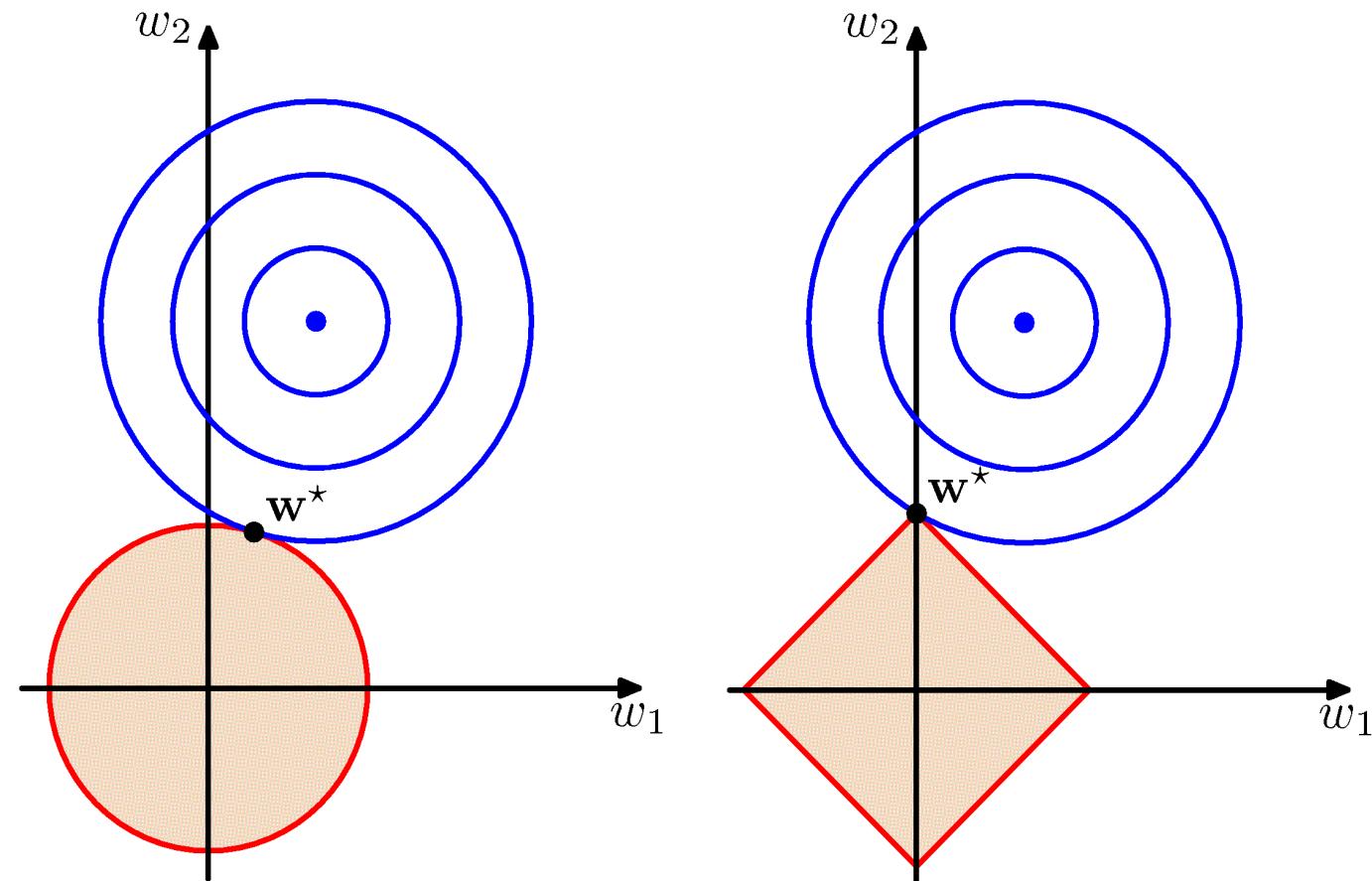
- Z bardziej uogólnionym regularyzatorem otrzymujemy

$$\frac{1}{2} \sum_{n=1}^N \{t_n - \mathbf{w}^T \phi(\mathbf{x}_n)\}^2 + \frac{\lambda}{2} \sum_{j=1}^M |w_j|^q$$



Regularyzowana MNK (3)

- Lasso ma tendencje do bardziej rozproszonych rozwiązań niż regularyzator kwadratowy.



Metody kernelowe

- Podejście w oparciu o funkcje bazowe daje możliwość poszerzenia przestrzenni cech czyniąc proste metody jak regresja liniowa znacznie pojemniejszymi
- Przykład:
 - Wejście: x
 - Cechy (funkcje bazowe) $x, x^2, x^3, \sin(x), \dots$
- Dwa potencjalne ograniczenia:
 - Wydajność obliczeniowa: jak znaleźć właściwe funkcje bazowe
 - Regularyzacja: jak uniknąć przetrenowania?
- Metody kernelowe próbują rozwiązać powyższe kwestie

Definicja jądra

- Założymy, że $\phi(\mathbf{x})$ mapuje D -wymiarowy wektor wejściowy \mathbf{x} na wielowymiarową (nieskończonie) przestrzeń cech
- Proste metody bazują na prostym iloczynie wektorów cech, $\phi(\mathbf{x}_1)^T \phi(\mathbf{x}_2)$
- Dla pewnych przestrzeni można wykorzystać “kernel trick” do wyznaczenia iloczynu $\phi(\mathbf{x}_1)^T \phi(\mathbf{x}_2)$ tylko przy użyciu wektora wejścia:

$$\phi(\mathbf{x}_1)^T \phi(\mathbf{x}_2) = k(\mathbf{x}_1, \mathbf{x}_2)$$

- $k(\mathbf{x}_1, \mathbf{x}_2)$ jest jądrem
- Sprawdzenie czy $k(x, y)$ jest właściwym jądrem zależy jedynie o właściwości samej funkcji jądrowej, bez konieczności weryfikacji macierzy cech

Przykłady funkcji jądrowych

- $k(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{x}_1^T \mathbf{x}_2$
- $k(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{x}_1^T \Sigma^{-1} \mathbf{x}_2$
(Σ^{-1} symetryczna dodatnio zdefiniowana)
- $k(\mathbf{x}_1, \mathbf{x}_2) = \exp(-\|\mathbf{x}_1 - \mathbf{x}_2\|^2 / 2\sigma^2)$
- $k(\mathbf{x}_1, \mathbf{x}_2) = \exp(-\frac{1}{2} \mathbf{x}_1^T \Sigma^{-1} \mathbf{x}_2)$
(Σ^{-1} symetryczna dodatnio zdefiniowana)
- $k(\mathbf{x}_1, \mathbf{x}_2) = p(\mathbf{x}_1)p(\mathbf{x}_2)$

Modularność

Metody kernelowe składają się z dwu części:

- 1) Wybór jądra (zadanie nietrywialne)
- 2) Właściwy algorytm go wykorzystujący

Modularność: możemy wykorzystać dowolny kernel z dowolnym algorytmem.

Przykładowe jądra:

$$k(x, y) = e^{(-\|x-y\|^2/c)}$$

$$k(x, y) = (\langle x, y \rangle + \theta)^d$$

$$k(x, y) = \tanh(\alpha \langle x, y \rangle + \theta)$$

$$k(x, y) = \frac{1}{\sqrt{\|x - y\|^2 + c^2}}$$

Przykładowe algorytmy:

- SVM (Support Vector Machine)
- Fisher discriminant analysis
- Regresja kernelowa
- kernel PCA
- kernel CCA

Proces Gaussowski

- Dla regresji liniowej: $\mathbf{y} = \Phi \mathbf{w}$
- Wykorzystując macierz Φ , wektor predykcji to $y(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x})$
- Jeśli (najprościej) \mathbf{w} : $p(\mathbf{w}) = \mathcal{N}(\mathbf{w} | \mathbf{0}, \alpha^{-1} \mathbf{I})$
- wtedy $\mathbb{E}[\mathbf{y}] = \Phi \mathbb{E}[\mathbf{w}] = \mathbf{0}$

$$\text{cov}[\mathbf{y}] = \mathbb{E} [\mathbf{y} \mathbf{y}^T] = \Phi \mathbb{E} [\mathbf{w} \mathbf{w}^T] \Phi^T = \frac{1}{\alpha} \Phi \Phi^T = \mathbf{K}$$

- **K** nazywamy macierzą Grama, gdzie

$$K_{nm} = k(\mathbf{x}_n, \mathbf{x}_m) = \frac{1}{\alpha} \phi(\mathbf{x}_n)^T \phi(\mathbf{x}_m)$$

- Tym samym: *korelacja dwu predykcji równa się jądru wyznaczonemu dla właściwych wejść*

$$\mathbb{E} [y(\mathbf{x}_n) y(\mathbf{x}_m)] = k(\mathbf{x}_n, \mathbf{x}_m)$$

Przykład: $k(x, x') = \exp(-\theta |x - x'|)$

Proces Gaussa: “uczenie” i predykcja

- Jak poprzednio zakładamy $p(t_n|y_n) = \mathcal{N}(t_n|y_n, \beta^{-1})$
- Wiarygodność wektora wyjściowego $p(\mathbf{t}|\mathbf{y}) = \mathcal{N}(\mathbf{t}|\mathbf{y}, \beta^{-1}\mathbf{I}_N)$
- Biorąc $p(\mathbf{y}) = \mathcal{N}(\mathbf{y}|\mathbf{0}, \mathbf{K})$, otrzymujemy graniczną dystrybucję predykcji wyjścia:

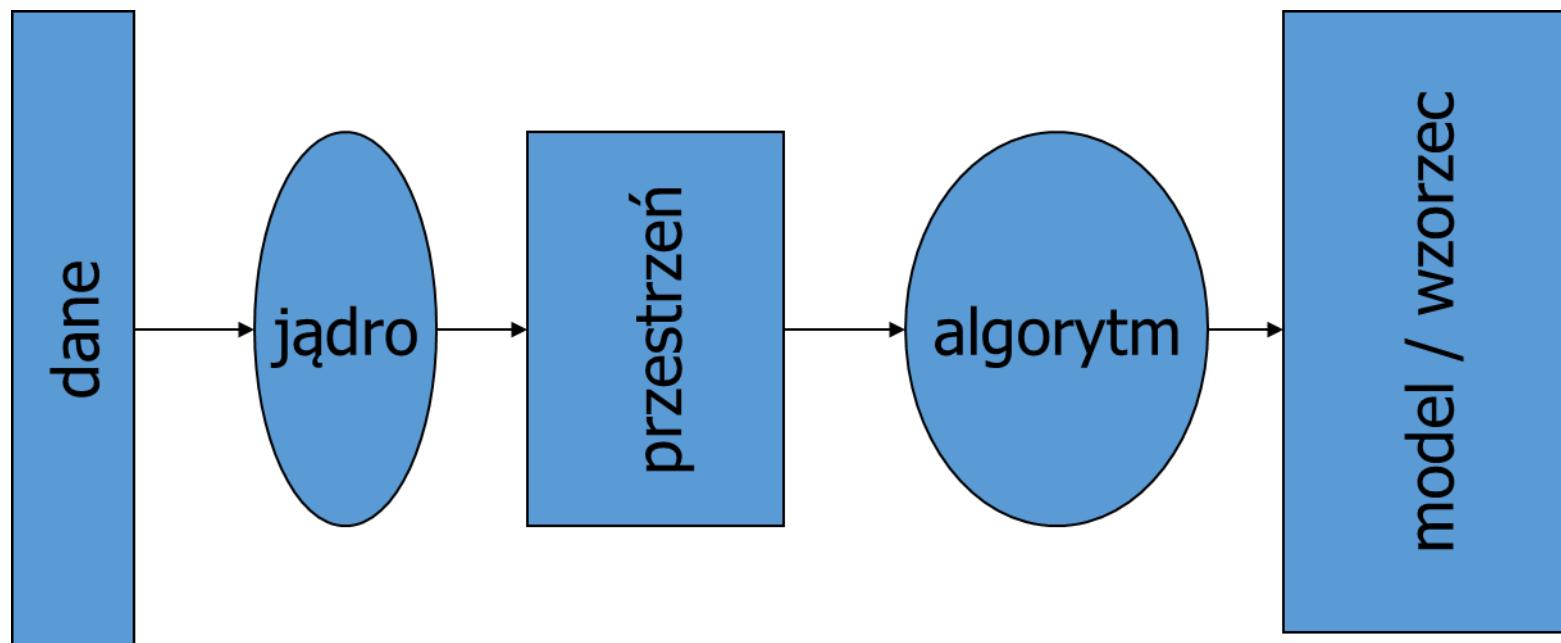
$$p(\mathbf{t}) = \int p(\mathbf{t}|\mathbf{y})p(\mathbf{y}) d\mathbf{y} = \mathcal{N}(\mathbf{t}|\mathbf{0}, \mathbf{C})$$

- gdzie $C(\mathbf{x}_n, \mathbf{x}_m) = k(\mathbf{x}_n, \mathbf{x}_m) + \beta^{-1}\delta_{nm}$
- Prognozy oparte są na: $p(\mathbf{t}_{N+1}) = \mathcal{N}(\mathbf{t}_{N+1}|\mathbf{0}, \mathbf{C}_{N+1})$
- gdzie $\mathbf{C}_{N+1} = \begin{pmatrix} \mathbf{C}_N & \mathbf{k} \\ \mathbf{k}^T & c \end{pmatrix} \quad \mathbf{k}_n = k(\mathbf{x}_n, \mathbf{x}_{N+1})$
 $c = k(\mathbf{x}_{N+1}, \mathbf{x}_{N+1}) + \beta^{-1}$
- $p(t_{N+1}|\mathbf{t})$ jest Gaussowski z:

$$m(\mathbf{x}_{N+1}) = \mathbf{k}^T \mathbf{C}_N^{-1} \mathbf{t}$$

$$\sigma^2(\mathbf{x}_{N+1}) = c - \mathbf{k}^T \mathbf{C}_N^{-1} \mathbf{k}$$

Metody jądrowe: plug and play



Wygładzanie (interpolacja) liniowa

- Wygładzanie jądrowe opisuje funkcje wagowe $W_{ni}(x)$ poprzez funkcję gęstości K wraz z parametrem skali, który modyfikuje (poprawia) rozmiar i formę wag w okolicy x .
- Jadro K jest gładką, ograniczoną i symetryczną funkcją rzeczywistą całkowaną do 1.
- Wagi są opisane jako:

$$W_{hi}(x) = K_h(x - X_i) / \hat{f}_h(x)$$

- gdzie

$$\hat{f}_h(x) = n^{-1} \sum_{i=1}^n K_h(x - X_i)$$

$$K_h(u) = h^{-1} K(u/h)$$

Estymata k-NN (k-Nearest Neighbor)

- Dla k-NN, otoczenie zdefiniowane jest poprzez te zmienne X, które są pomiędzy k-najbliższymi sąsiadami x względem odległości euklidesowej.
- Estymator k-NN definiujemy jako:

$$\hat{m}_k(x) = n^{-1} \sum_{i=1}^n W_{ki}(x) Y_i$$

- gdzie $\{W_{ki}(x)\}_{i=1, \dots, n}$ definiujemy jako zbiór indeksów

$J_x = \{i : X_i \text{ } i \text{ jest jedną z } k \text{ najbliższych obserwacji } x\}$

- i
- $$W_{ki}(x) = \begin{cases} n/k, & \text{if } i \in J_x \\ 0 & \text{inne} \end{cases}$$

Estymata k-NN

- Parametr wygładzania k jest odpowiedzialny za gładkość estymaty (pasmo wygładzania).
- Jeśli $k > n$, estymata k -NN odpowiada średniej.
- Jeśli $k = 1$, obserwacje są odtworzone w X_i i dla x pomiędzy dwoma sąsiednimi predykowanymi zmiennymi jest odtworzone jako skok pomiędzy nimi.

Jednowymiarowe wygładzanie jądrowe

- k-NN: $\hat{f}(x) = Ave(y_i \mid x_i \in N_k(x))$
- 30-NN nie jest gładka, ponieważ $\hat{f}(x)$ jest nieciągła w x.
- Dyskretne zmiany powodują nieciągłość $\hat{f}(x)$.

Jednowymiarowe wygładzanie jądrowe

- Nadaraya-Watson Kernel średnia ważona:

$$\hat{f}(x_0) = \frac{\sum_{i=1}^N K_\lambda(x_0, x_i) y_i}{\sum_{i=1}^N K_\lambda(x_0, x_i)}$$

- Jądro kwadratowe Epanechnikova:

$$K_\lambda(x_0, x) = D \left(\frac{|x - x_0|}{\lambda} \right)$$

Lokalna regresja liniowa

- Kwestie graniczne

- Złe uwarunkowanie na ograniczeniach ze względu na niesymetrię w okolicy
- Przesunięcie jest usunięte przez dopasowanie liniowe

Lokalna regresja liniowa

- Lokalna ważona regresja liniowa realizuje korekcję pierwszego rzędu
- Oddzielne ważone MNK dla każdego punktu x_0 :

$$\min_{\alpha(x_0), \beta(x_0)} \sum_{i=1}^N K_\lambda(x_0, x_i) [y_i - \alpha(x_0) - \beta(x_0)x_i]^2$$

- aby estymować: $\hat{f}(x_0) = \hat{\alpha}(x_0) + \hat{\beta}(x_0)x_0$
- $b(x)^T = (1, x)$; macierz regresji B : $N \times 2$ oraz i -tym wierszem $b(x)^T$;

$$W_{N \times N}(x_0) = \text{diag}(K_\lambda(x_0, x_i)), i = 1, \dots, N$$

$$\hat{f}(x_0) = b(x_0)^T (B^T W(x_0) B)^{-1} B^T W(x_0) y = \sum_{i=1}^N l_i(x_0) y_i$$

Lokalna regresja liniowa

- Wagi $l_i(x_0)$ łączą lokalne wygładzanie jądrowe $K_\lambda(x_0, \cdot)$ i metodę najmniejszych kwadratów – **Equivalent Kernel**

Lokalna regresja wielomianowa

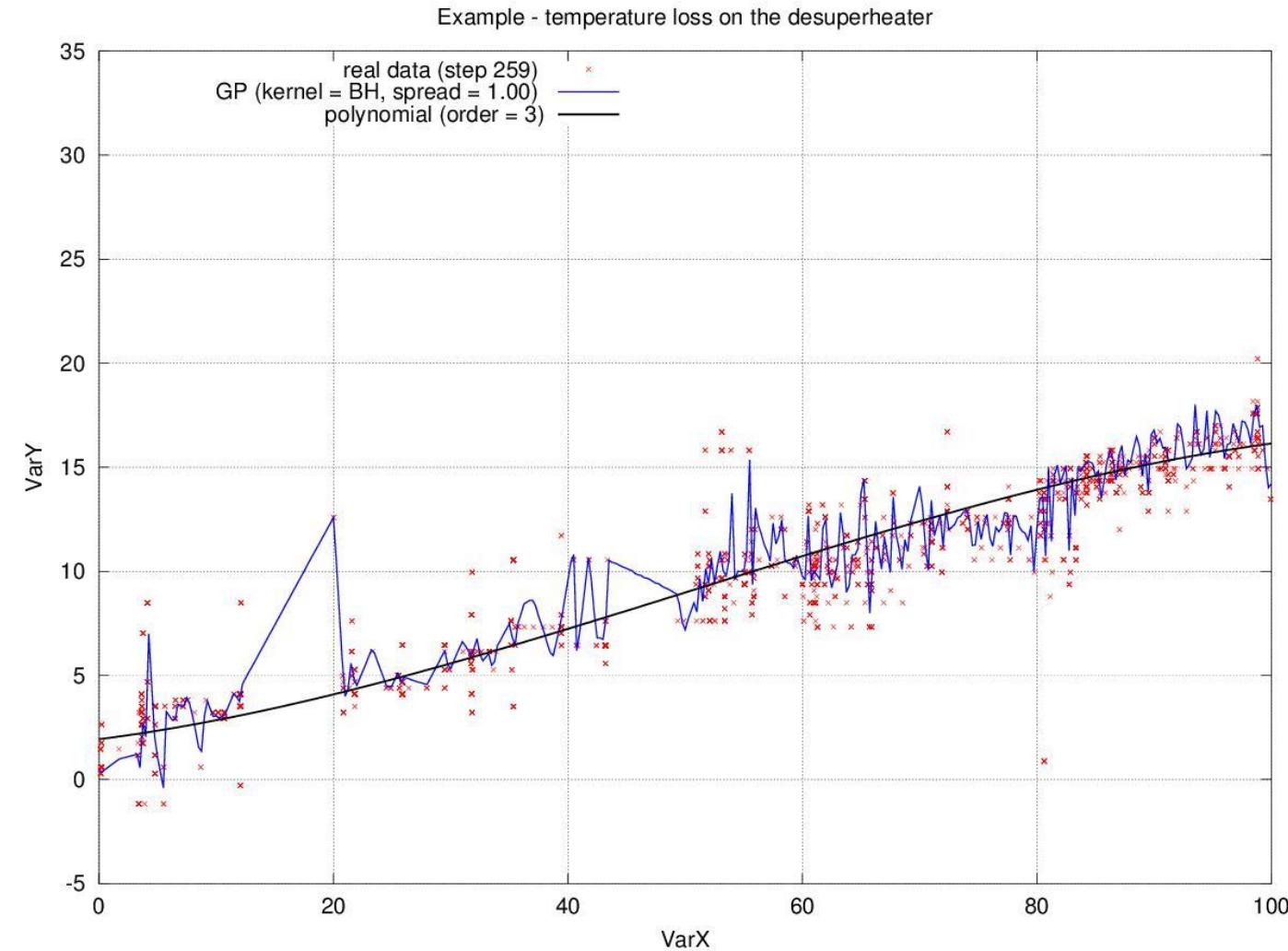
- Lokalne dopasowanie wielomianu o stopniu d

Ograniczenia stałych funkcji jądrowych

- Rozmieszczać równomiernie – przekleństwo wymiarowości.
- Może grupowanie danych
 - Metoda k-means
 - Metoda C-means
 - Sieci RBF
 - ...

Przykład #3 (1) - zawór

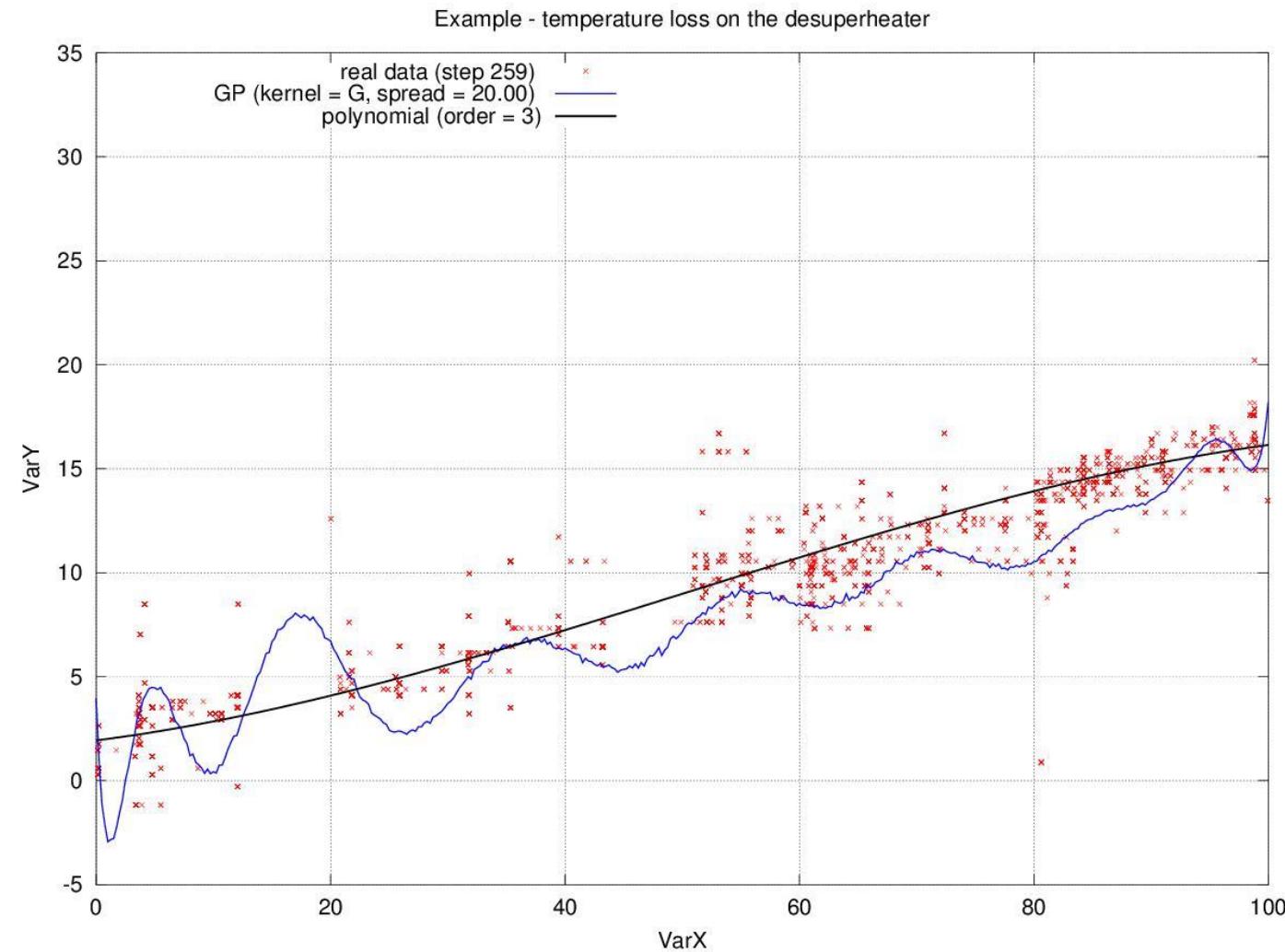
Biharmonic kernel



Przykład #3 (2) - zawór

Gaussian kernel

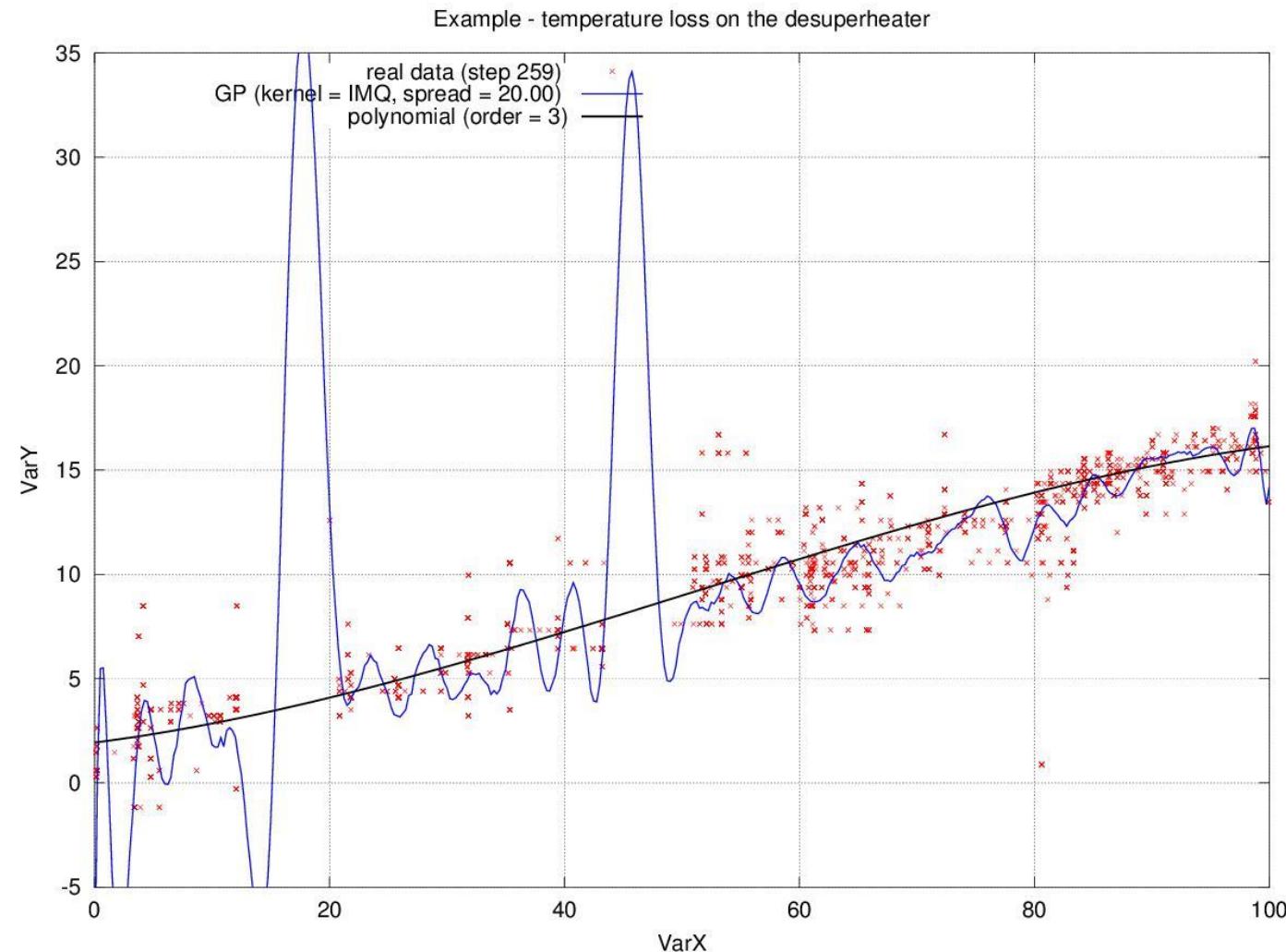
$$\phi(r) = e^{-(\epsilon r)^2}$$



Przykład #3 (3) - zawór

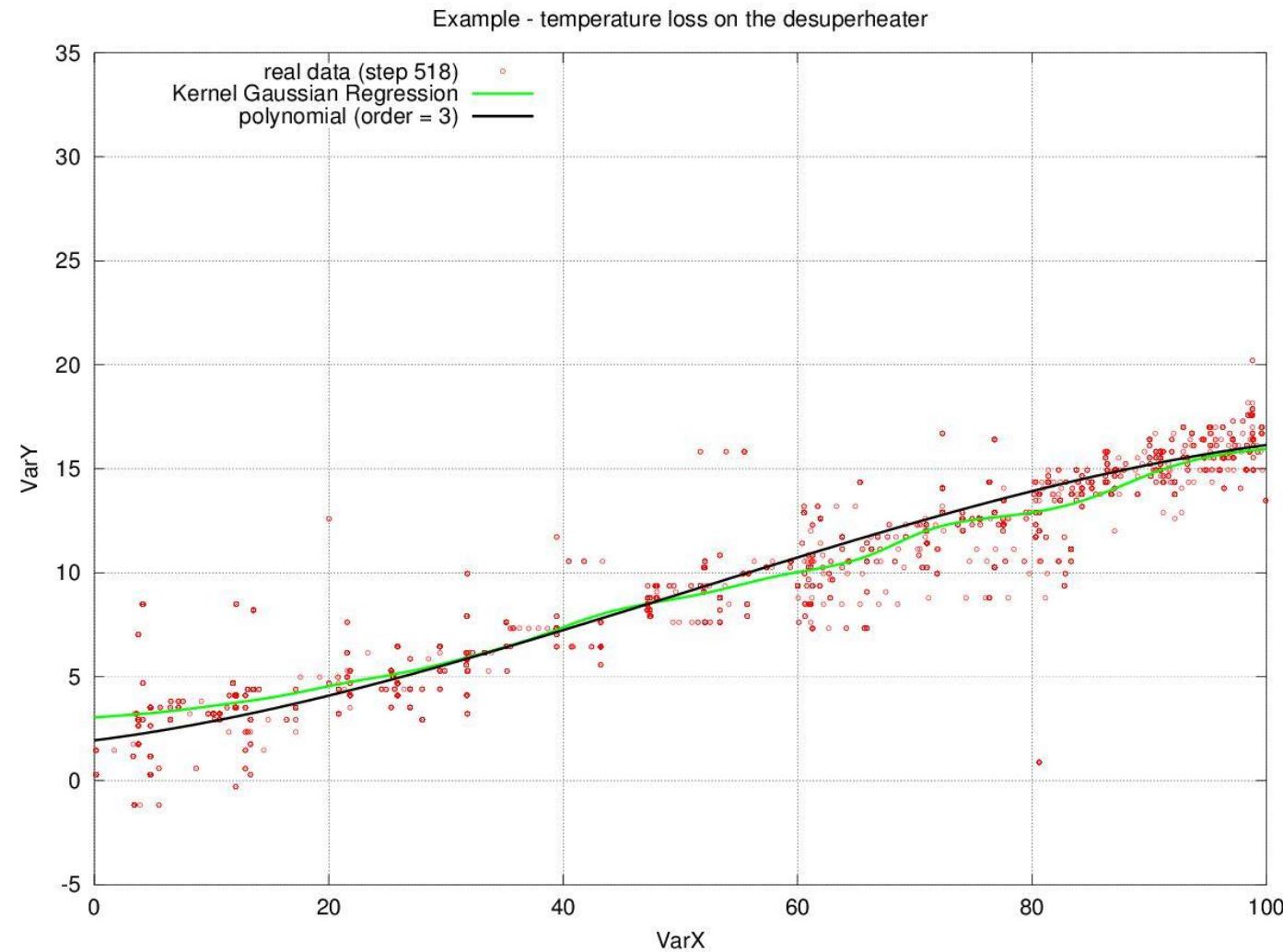
Inverse
Multiquadric
kernel

$$\phi(r) = \frac{1}{\sqrt{1 + (\varepsilon r)^2}}$$



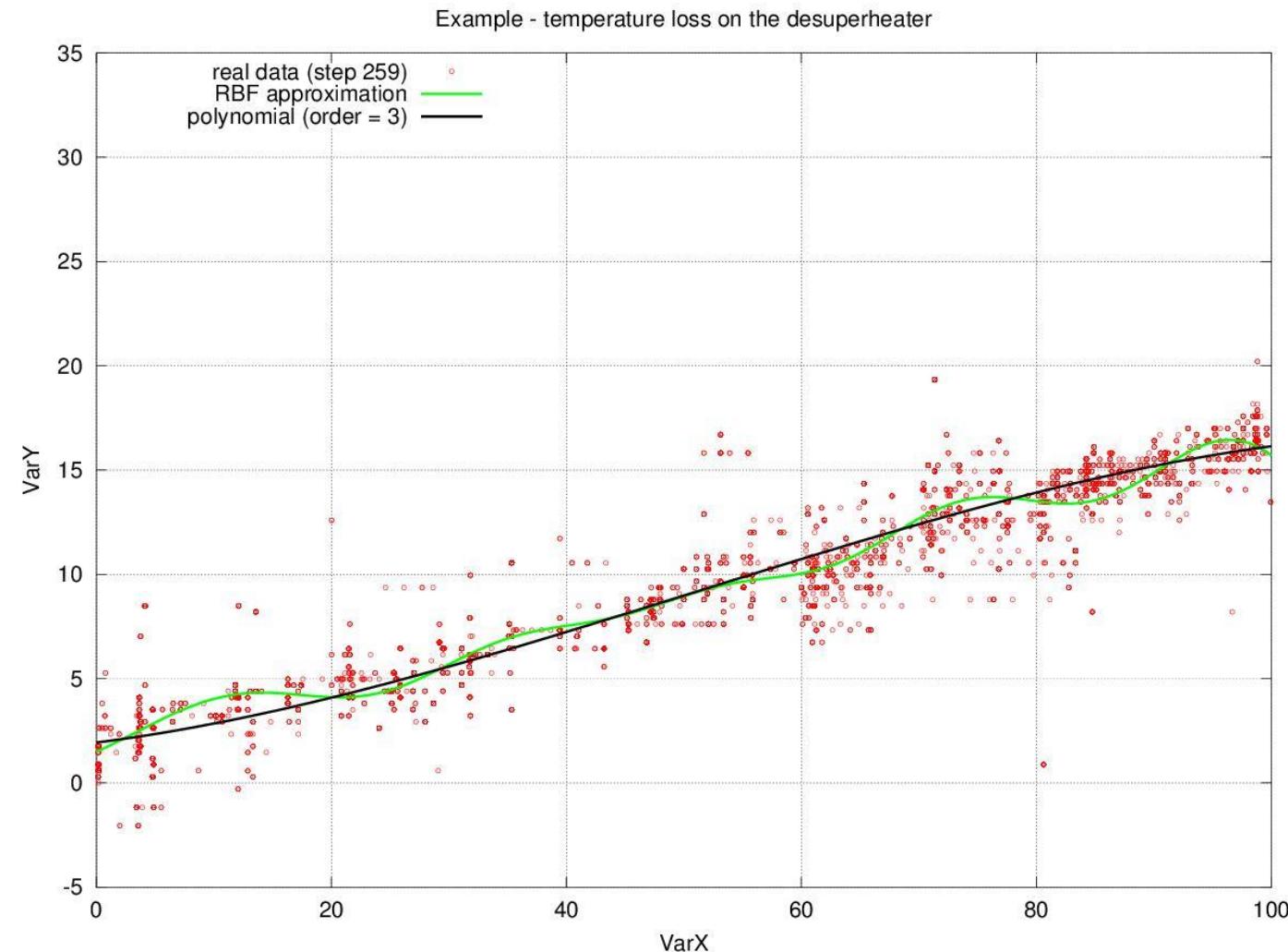
Przykład #3 (4) - zawór

KGR



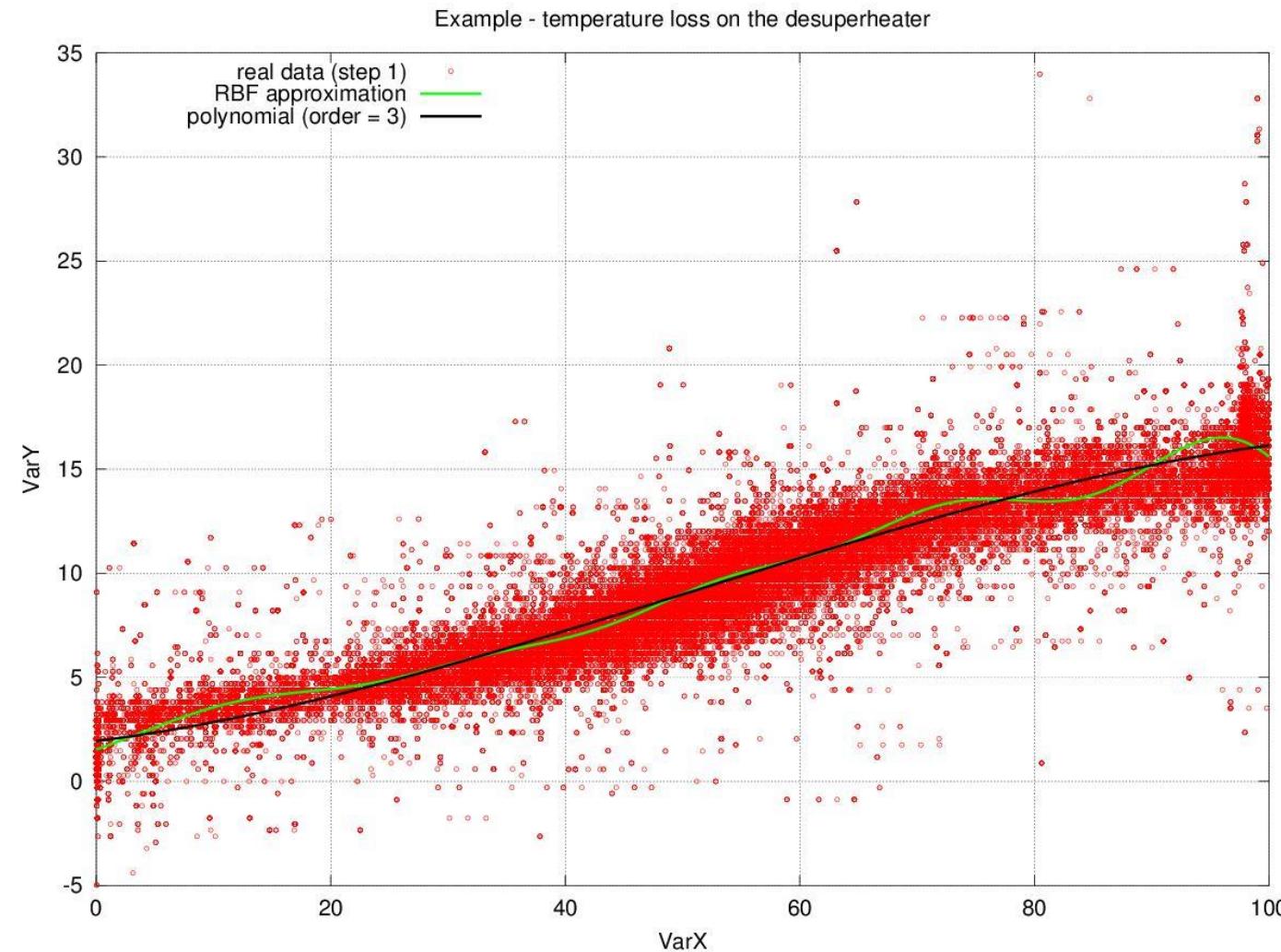
Przykład #3 (5) - zawór

Radial Basis
Function
multiple-running



Przykład #3 (6) - zawór

Radial Basis
Function
single shot





MI

Metody Identyfikacji

wykład #10a

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

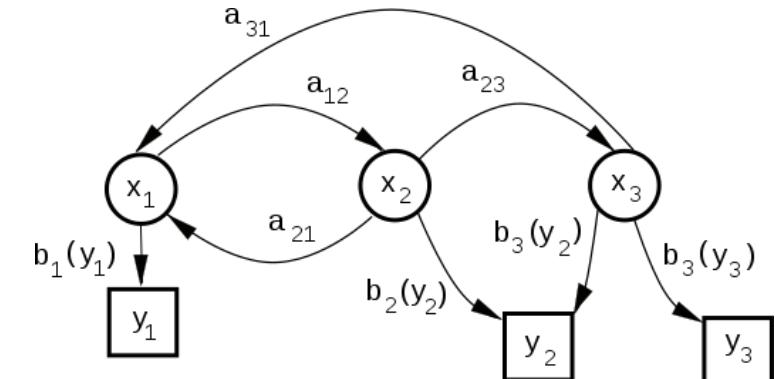
Estymacja stanu

- Zadanie polega na estymacji stanu procesu $\hat{x}(k)$ na podstawie wejścia $u(l)$ i wyjścia $y(l)$ 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



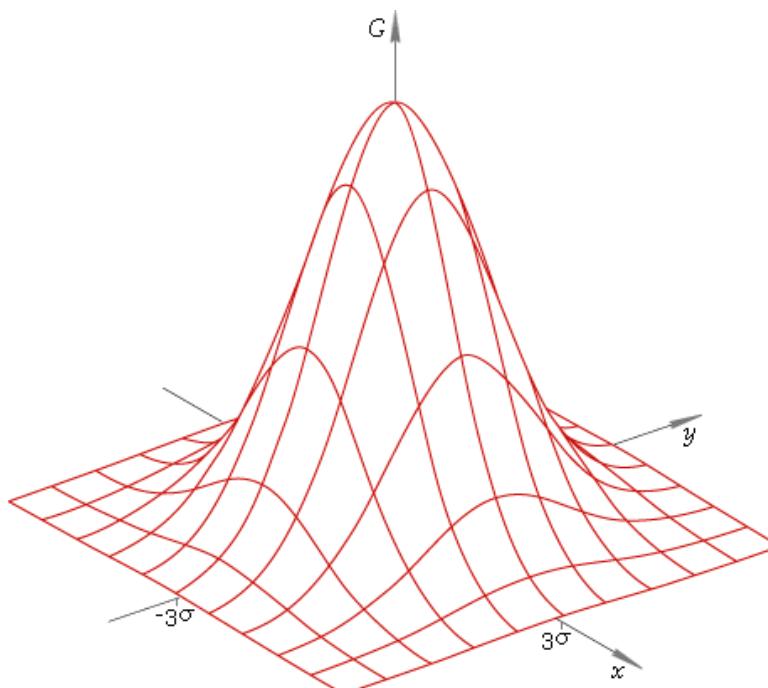
- Filtry Kalmana

- 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

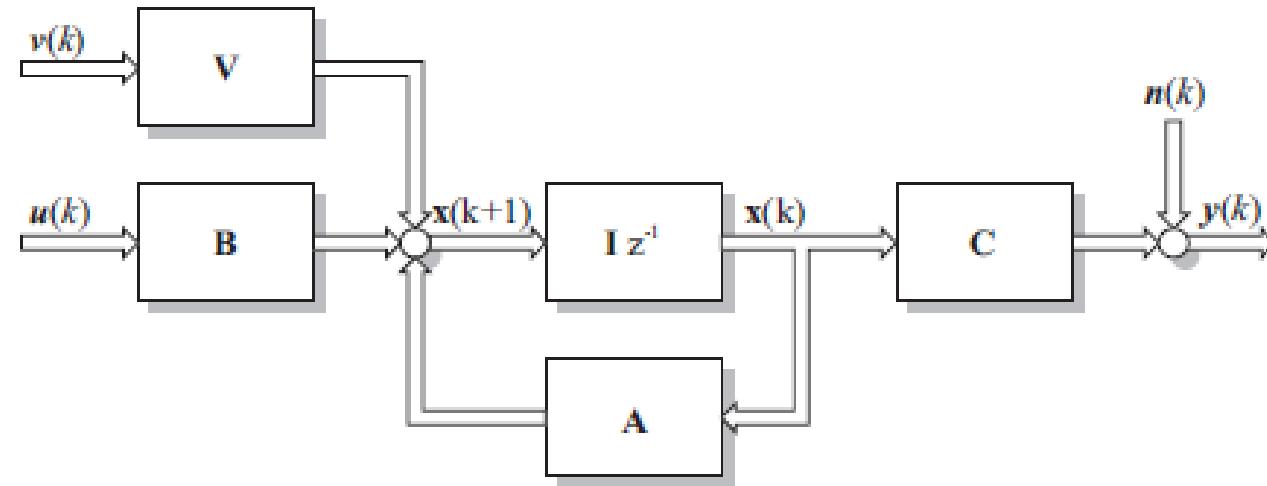


Filtr Kalmana

- 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



$$\begin{aligned}x(k+1) &= A \cdot x(k) + B \cdot u(k) + V \cdot v(k) \\y(k) &= C \cdot x(k) + n(k)\end{aligned}$$

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

$$\begin{aligned}E\{v(k)v^T(k)\} &= M \\E\{n(k)n^T(k)\} &= N\end{aligned}$$

Filtr Kalmana

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

$$\begin{aligned} V &= \mathbb{E}\left\{\|\hat{x}(k+1) - x(k+1)\|_2^2\right\} \\ &= \mathbb{E}\left\{(\hat{x}(k+1) - x(k+1))^T(\hat{x}(k+1) - x(k+1))\right\} \end{aligned}$$

- 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

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

Filtr Kalmana – predykcja

- Wyznaczamy predykcję

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

- Zakłόcenie $v(k)$ nie jest znane.
- Ma zerową wartość oczekiwana.
- 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 $P^-(k+1)$

$$\begin{aligned}
 P^-(k+1) &= E\left\{(\hat{x}(k+1|k) - x(k+1))(\hat{x}(k+1|k) - x(k+1))^T\right\} \\
 &= E\left\{(A\hat{x}(k) - Ax(k) + Vv(k))(A\hat{x}(k) - Ax(k) + Vv(k))^T\right\} \\
 &= AE\left\{(\hat{x}(k) - x(k))(\hat{x}(k) - x(k))^T\right\}A^T \\
 &\quad + AE\left\{(\hat{x}(k) - x(k))v^T\right\}V^T \\
 &\quad + VE\left\{v(k)(\hat{x}(k) - x(k))^T\right\}A^T \\
 &\quad + VE\left\{v(k)v^T(k)\right\}V^T
 \end{aligned}$$

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):

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

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

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

- Co prowadzi do postaci macierzy kowariancji

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

Filtr Kalmana – korekcja

- optymalne wzmacnienie filtru Kalmana

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

- poprawiony wektor stanu

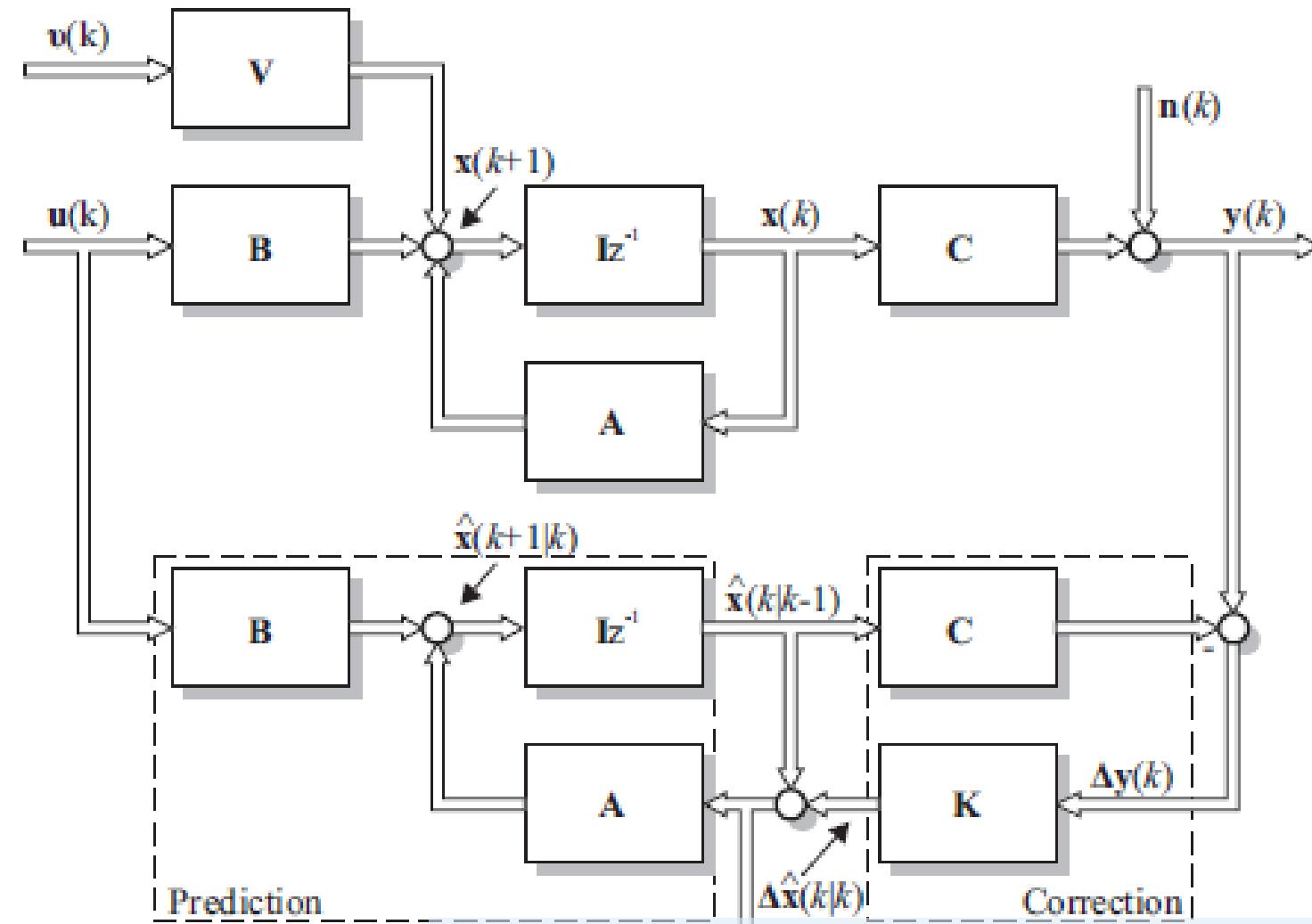
$$\hat{x}(k + 1|k + 1) = \hat{x}(k + 1|k) + K(k + 1) \cdot (y(k + 1) - C \cdot \hat{x}(k + 1|k))$$

- poprawiona macierz kowariancji

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

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

Filtr Kalmana – schemat blokowy



optymalna estymata stanu na jeden krok naprzód

Niezmienność

- 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 $k \rightarrow \infty$
- 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

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

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

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 filtra 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 losową stałą \rightarrow "voltage" odczytywaną ze źródła
- Posiada ona stałą wartość aV [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 I
- Zakładamy jednostkową początkową kowariancję $P_0 = I$
a początkowy stan zerowy $x_0 = 0$.

Prosty przykład – cd.

$$\begin{aligned}x_k &= Ax_{k-1} + Bu_k + w_k \\&= x_{k-1} + w_k\end{aligned}$$

$$\begin{aligned}z_k &= Hx_k + v_k \\&= x_k + v_k\end{aligned}$$

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

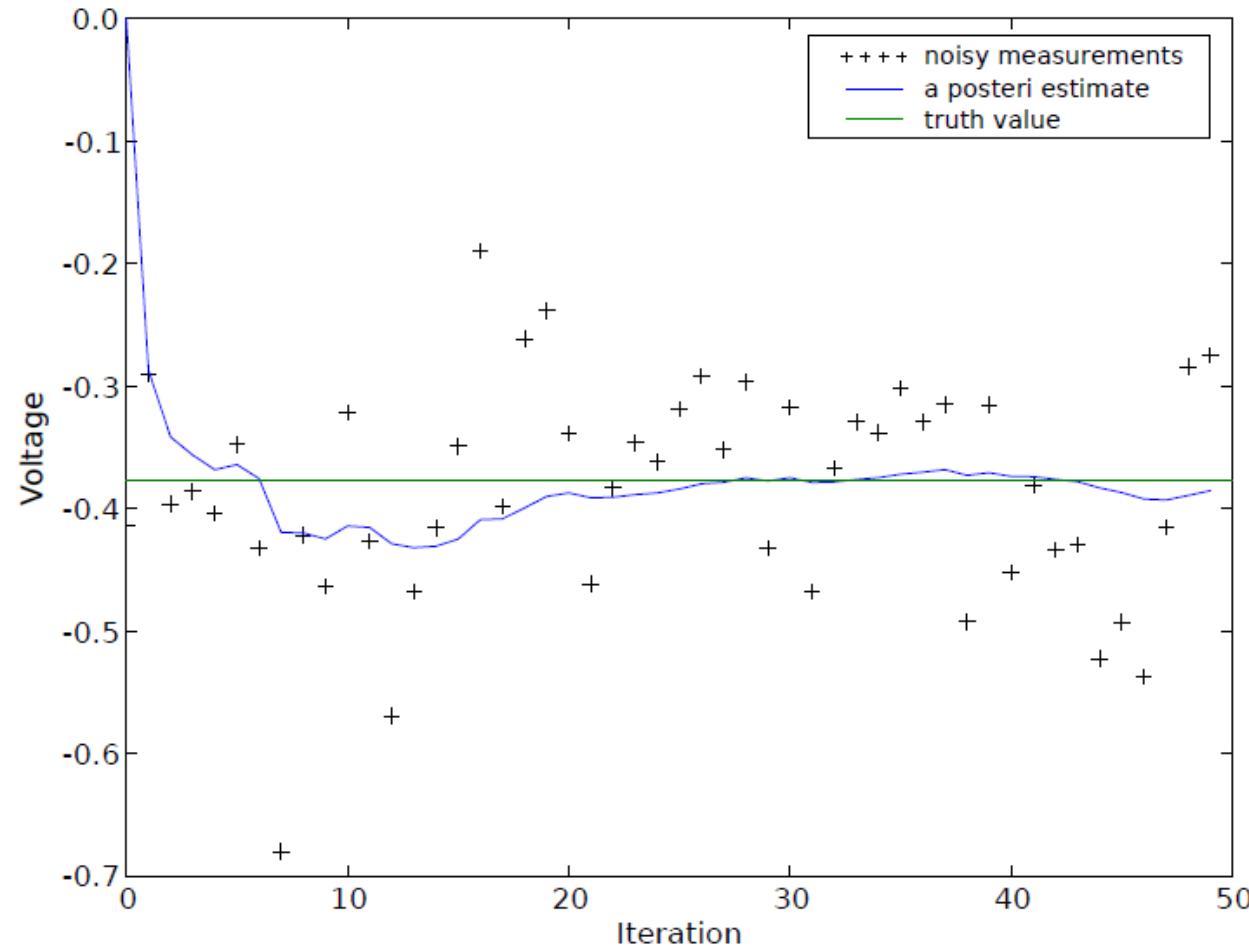
Prosty przykład – cd.

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

Prosty przykład – cd.

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	$\hat{x}_k^- = \hat{x}_{k-1} = 0$ $P_k^- = P_{k-1} = 1$	$\hat{x}_k^- = 0.355$ $P_k^- = 0.091$								
Measurement Update	$K_k = 1 / (1 \cdot 0.1)$ = 0.909 $\hat{x}_k = 0.909 (0.390 - 0)$ = 0.35 $P_k = (1 - 0.909) \cdot 1$ = 0.091	$K_k = 0.091 / (0.091 \cdot 0.1)$ = 0.476 $\hat{x}_k = 0.355 \cdot 0.476 (0.500 - 0.355)$ = 0.424 $P_k = (1 - 0.476) \cdot 0.091$ = 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 wzmacnień
- Inicjalizacja
- Kwestia brakujących danych
- Testowanie
- Metody obliczeniowe

Zastosowania filtru Kalmana

- Popularne zastosowania
 - nawigacja, naprowadzanie, śledzenie radarowe, przeszukiwanie sonarowe, wyznaczanie orbit satelitów, 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

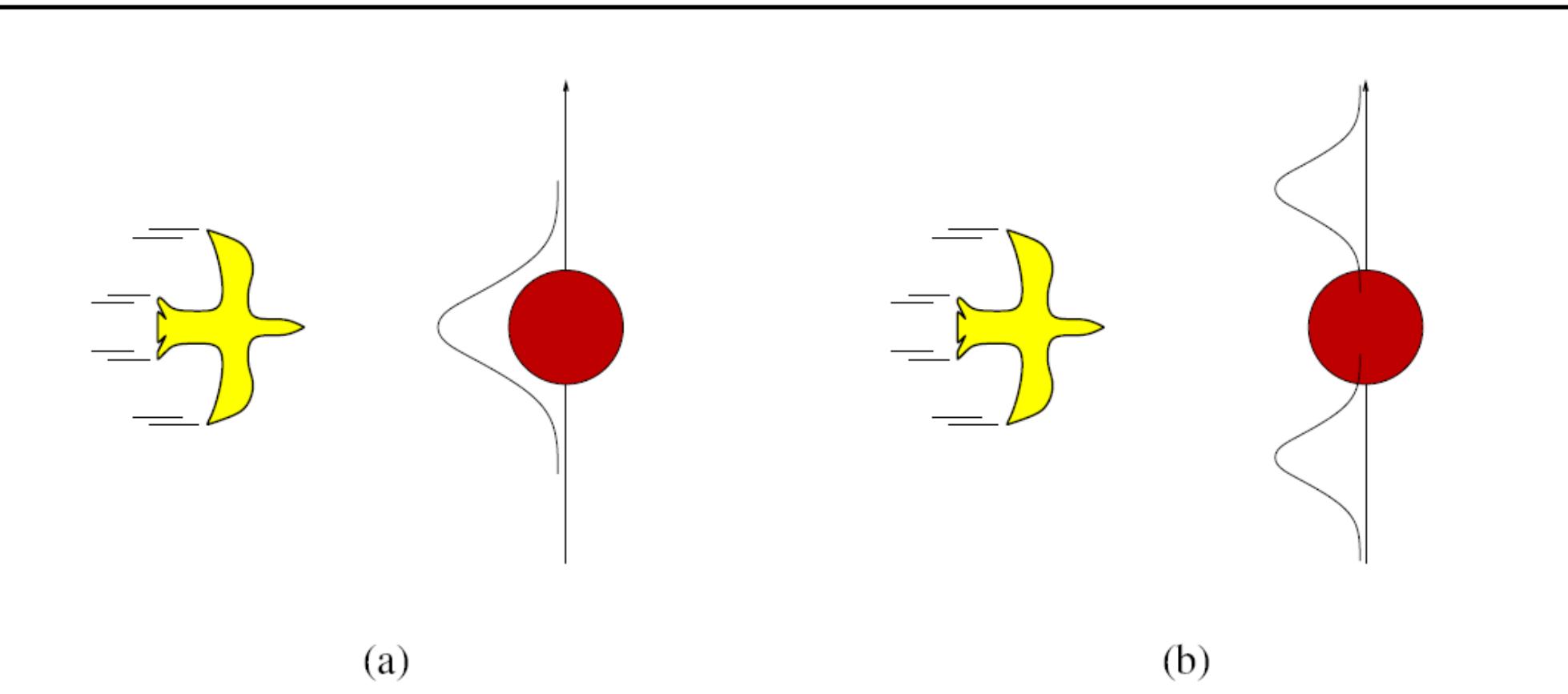


Figure 15.10 A bird flying toward a tree (top views). (a) A Kalman filter will predict the location of the bird using a single Gaussian centered on the obstacle. (b) A more realistic model allows for the bird's evasive action, predicting that it will fly to one side or the other.

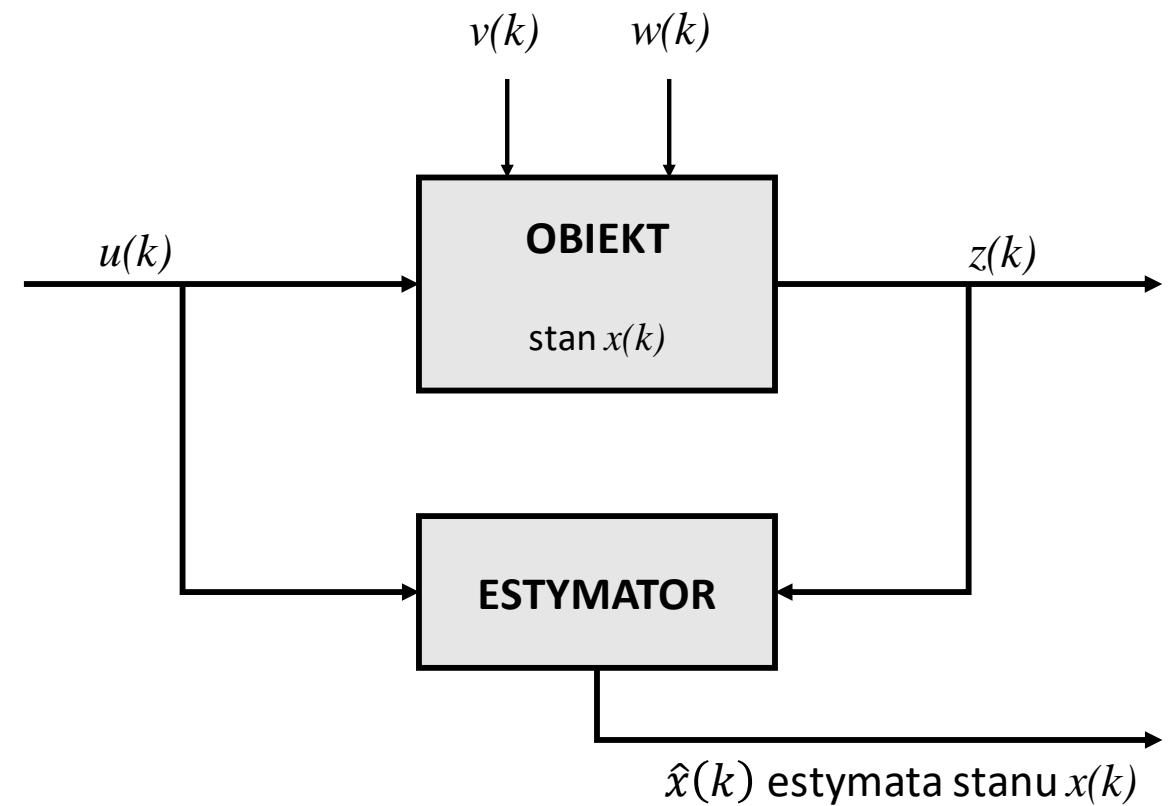
Obserwator stanu

- 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 $x(o)$, $\{v(\cdot)\}$, $\{w(\cdot)\}$ mają opis stochastyczny



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 $\{u(\cdot)\}$, $\{z(\cdot)\}$

- 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.

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 $B_1, B_2, \dots B_n$, 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)

Twierdzenie Bayesa – przypadek 1

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

Twierdzenie Bayesa – przypadek 2

- Przypadek 2: Rodzina ma dwoje dzieci. Jeśli jedna to dziewczynka, to jakie jest prawdopodobieństwo, że druga jest dziewczynką?
- Odpowiedź:

• Oznaczamy: $a = \begin{cases} 1: & \text{nie ma chłopców} \\ 0: & \text{jest chłopiec} \end{cases}$ $b = \begin{cases} 1: & \text{nie ma dziewczynek} \\ 0: & \text{jest dziewczynka} \end{cases}$

• 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:

$$p_{a|b}(1|0) = \frac{p_a(1) \cdot p_b|_a(0|1)}{p_b(0)} = \frac{\frac{1}{4} \cdot \frac{1}{4}}{\frac{3}{4}} = \frac{1}{12}$$

Twierdzenie Bayesa – przypadek 3

- Przypadek 3: 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: $a = \begin{cases} 1: & \text{pacjent nie ma raka} \\ 0: & \text{pacjent ma raka} \end{cases}$ $b = \begin{cases} 1: & \text{mammograf daje wynik negatywny} \\ 0: & \text{mammograf daje wynik pozytywny} \end{cases}$

- Zatem szukamy:

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

Twierdzenie Bayesa – przypadek 3

- Wyznaczamy:
 - $p_{b|a}(0|0) = 0.9$ (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: B_1, \dots, B_N
- Można założyć niezależność, tzn. $p(B_1, \dots, B_N | A) = p(B_1 | A) \dots p(B_N | A)$
Założenie, że szum jest niezależny
- Wtedy

$$p(A | B_1, \dots, B_N) = \frac{p(A) \cdot \prod_i p(B_i | A)}{p(B_1, \dots, B_N)}$$

a priori *wiarygodność
obserwacji*
a posteriori *normalizacja*

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 ($k = 1, 2, \dots$)
$$x(k) = q_{k-1}(x(k-1), v(k-1))$$
$$z(k) = h_k(x(k), w(k))$$
- gdzie $x(0), \{v(\cdot)\}, \{w(\cdot)\}$ są niezależne o nieznanych funkcjach gęstości prawdopodobieństwa

Śledzenie Bayesowskie – cel

- Niech $z(1:k)$ oznacza zbiór $\{z(1), \dots, z(k)\}$. 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 $v(k-1)$
- $z(k-1) = h_{k-1}(x(k-1), w(k-1))$
- $z(k-2) = h_{k-2}(x(k-2), w(k-2))$, $x(k-2) = q_{k-3}(x(k-3), v(k-3))$, 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*

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

- $z(k)$ 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:

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

Algorytm rekurencyjny – podsumowanie

- Krok 1: PREDYKCJA

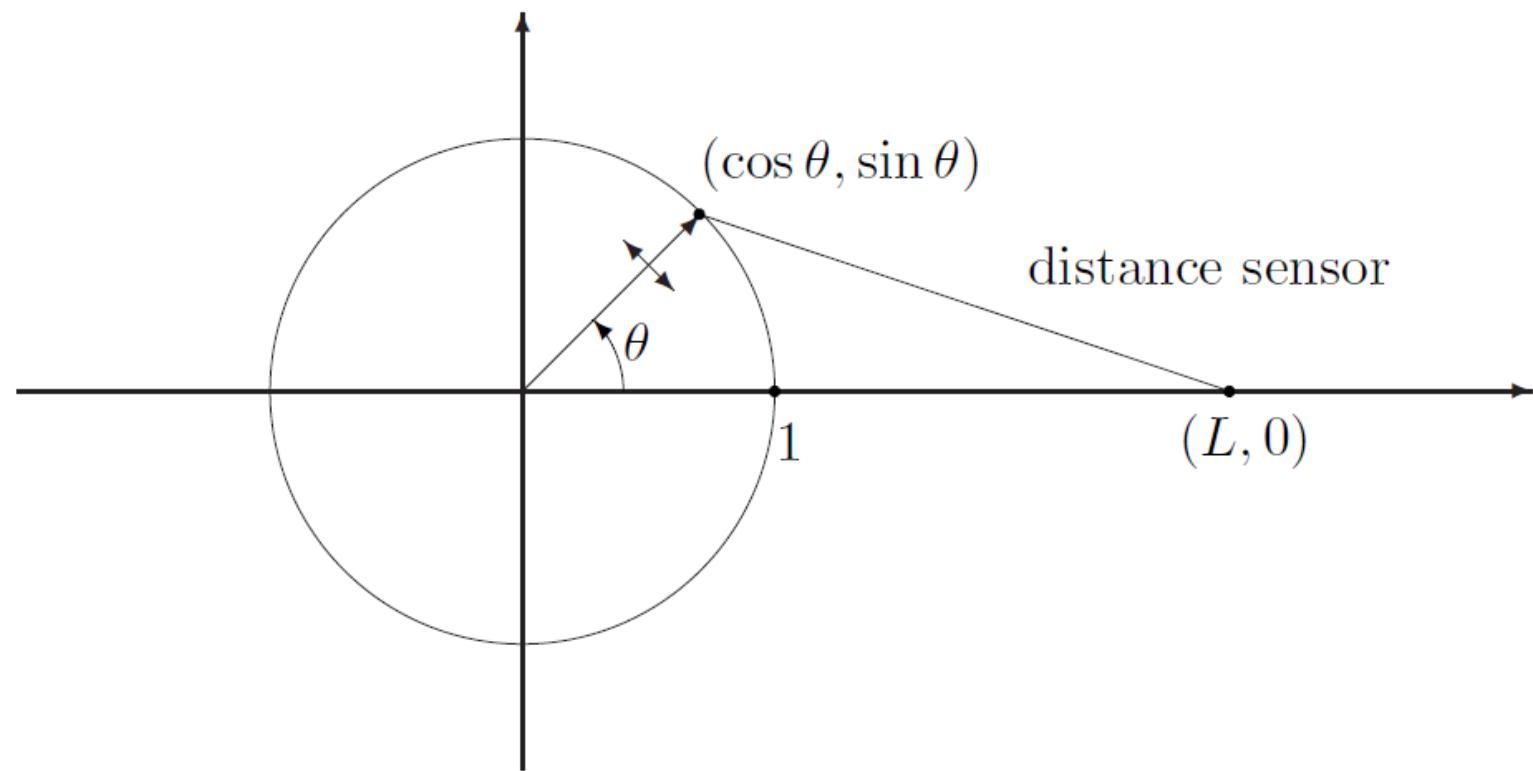
$$f(x(k)|z(1:k-1)) = \sum_{x(k-1) \in \mathcal{X}} \overbrace{f(x(k)|x(k-1))}^{\text{model}} \overbrace{f(x(k-1)|z(1:k-1))}^{\text{poprzednia iteracja}}, \quad k = 1, 2, \dots$$

- KROK 2: AKTUALIZACJA

$$f(x(k)|z(1:k)) = \frac{\overbrace{f(z(k)|x(k))}^{\text{model pomiaru}} \overbrace{f(x(k)|z(1:k-1))}^{a \text{ priori}}}{\underbrace{\sum_{\bar{x}(k) \in \mathcal{X}} f(z(k)|\bar{x}(k)) f(\bar{x}(k)|z(1:k-1))}_{\text{normalizacja}}}$$

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:

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

gdzie $\text{mod}(\cdot, N)$ oznacza modulo N operacji, tzn. $\text{mod}(N, N) = 0$ a $\text{mod}(-1, N) = N-1$.

- 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:

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

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

- Inicjalizacja: $f(x(0)) = 1/N$

Filtr Kalmana jako obserwator stanu

- Założmy, że obiekt jest stacjonarny:

$$x(k) = Ax(k-1) + u(k-1) + v(k-1)$$

$$z(k) = Hx(k) + w(k)$$

$$x(0) \sim \mathcal{N}(x_0, P_0), v(k-1) \sim \mathcal{N}(0, Q)$$

$$w(k) \sim \mathcal{N}(0, R)$$

- 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 $k \rightarrow \infty$.
- 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 filtra Kalmana wystarczająco długo otrzymamy stacjonarny rozkład błędów estymacji

Wykrywalność

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

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

- Idea wykrywalności:
 - Wykrywalność jest słabsza niż obserwowlność, tzn. (A, H) obserwowlne $\rightarrow (A, H)$ wykrywalne
 - Jeśli system jest wykrywalny, to wtedy:

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

Filtr Kalmana w stanie ustalonym

- Jeśli wariancja filtru Kalmana jest zbieżna, to wtedy:

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

- 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*)

Filtr Kalmana w stanie ustalonym

- 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 wzmacnienie w stanie ustalonym dane jest wzorem:

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

Filtr Kalmana w stanie ustalonym

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

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

liniowy stacjonarny system.

- Błąd estymacji:

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

bo inaczej nie zbiega

Filtr Kalmana w stanie ustalonym

- Co może się nie udać:

- $P_P(k)$ nie zbiega jak $k \rightarrow \infty$,
- $P_P(k)$ 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 ...

Rozszerzony Filtr Kalmana – Extended Kalman Filter

- Rozważmy nieliniowy system dyskretny w czasie:

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

$$\mathbf{E}[x(0)] = x_0, \text{Var}[x(0)] = P_0$$

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

$$\mathbf{E}[v(k-1)] = 0, \text{Var}[v(k-1)] = Q(k-1)$$

$$\mathbf{E}[w(k)] = 0, \text{Var}[w(k)] = R(k)$$

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

Rozszerzony Filtr Kalmana – Extended Kalman Filter (EKF)

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

Rozszerzony Filtr Kalmana – Extended Kalman Filter (EKF)

- Założymy, że wyznaczono $\hat{x}_m(k-1)$ oraz $P_m(k-1)$ jako aproksymację średniej oraz wariancji stanu $x(k-1)$ na podstawie pomiarów $x(1:k-1)$.
- Lineryzacja równania stanu wokół punktu $x(k-1) = \hat{x}_m(k-1)$ oraz $v(k-1) = E[v(k-1)] = 0$ prowadzi do:

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

$$+ \underbrace{\frac{\partial q_{k-1}(\hat{x}_m(k-1), u(k-1), 0)}{\partial x} \cdot (x(k-1) - \hat{x}_m(k-1))}_{=:A(k-1)} + \underbrace{\frac{\partial q_{k-1}(\hat{x}_m(k-1), u(k-1), 0)}{\partial v} \cdot v(k-1)}_{=:L(k-1)}$$

$$= A(k-1)x(k-1) + \underbrace{L(k-1)v(k-1)}_{=: \tilde{v}(k-1)} + \underbrace{q_{k-1}(\hat{x}_m(k-1), u(k-1), 0) - A(k-1)\hat{x}_m(k-1)}_{=: \xi(k-1)}$$

$$= A(k-1)x(k-1) + \tilde{v}(k-1) + \xi(k-1),$$

szum procesowy znane wejście

Rozszerzony Filtr Kalmana – Extended Kalman Filter (EKF)

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

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

- I otrzymujemy

$$\begin{aligned}\hat{x}_p(k) &= A(k-1)\hat{x}_m(k-1) + \xi(k-1) \\ &= q_{k-1}(\hat{x}_m(k-1), u(k-1), 0) \quad \text{wstawiając } \xi(k-1)\end{aligned}$$

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

Rozszerzony Filtr Kalmana – Extended Kalman Filter (EKF)

- Linearyzacja prowadzi do

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

↗ dodatkowy element

- $\tilde{w}(k)$ ma zerową wartość średnią oraz wariancję:

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

Rozszerzony Filtr Kalmana – Extended Kalman Filter (EKF)

- Zastosowanie aktualizacji filtru Kalmana do równania zlinearyzowanego prowadzi do:

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

$$\begin{aligned}\hat{x}_m(k) &= \hat{x}_p(k) + K(k) (z(k) - H(k)\hat{x}_p(k) - \zeta(k)) \\ &= \hat{x}_p(k) + K(k) (z(k) - h_k(\hat{x}_p(k), 0)) \quad \text{wstawiając } \zeta(k-1)\end{aligned}$$

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

Rozszerzony Filtr Kalmana – Extended Kalman Filter (EKF)

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

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

Rozszerzony Filtr Kalmana – Extended Kalman Filter (EKF)

1. Aktualizacja *a posteriori*

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

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

$$P_m(k) = (I - K(k) H(k)) P_p(k)$$

gdzie

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

Rozszerzony Filtr Kalmana – Extended Kalman Filter (EKF)

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

$$z[k] = h_k(x[k], w[k]) \quad \text{E}[w[k]] = 0, \text{Var}[w[k]] = R,$$
- Oznaczamy dyskretyzację:
$$x[k] := x(kT), T - \text{okres próbkowania.}$$
- 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.



MI
Metody Identyfikacji

wykład #10b

1. *Particle filter*

Filtr cząstek – Particle Filter

- Rozszerzenie Bayesowskiego estymatora stanu na **ogólny model nieliniowy** oraz **ogólne zakłócenia**.
- Idea:
 - Aproxymacja gęstości prawdopodobieństwa stanu za pomocą dużej ilości próbek, nazywanych cząstkami (*particles*).
 - W regionach o dużych wartościach funkcji gęstości jest dużo owych cząstek.
 - Rekurencyjny algorytm aktualizacji działa następująco:
 - i. Cząstki są propagowane w obiekcie
 - ii. Następnie są ważone w zależności od wiarygodności pomiaru
 - iii. Podczas nowego kroku próbkowania wybieramy nowy zbiór cząstek, też o równych wagach.

Filtr cząstek – Particle Filter

- Rozważamy ogólny model nieliniowy:

$$x(k) = q_{k-1}(x(k-1), v(k-1)), \quad k = 1, 2, \dots$$

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

- Ogólnie wyprowadzenie będzie dla systemu bez wejścia $u(k-1)$, jednakże można dyskusje rozszerzyć na system zawierający wejście.

Filtr cząstek – Particle Filter

- Próbkowanie Monte Carlo
- Wykorzystanie dużej liczby próbek w celu aproksymacji funkcji gęstości prawdopodobieństwa
- Podstawa algorytmu filtru cząstek

Filtr cząstek – Particle Filter

- Rozważmy dyskretną zmienną losową y
 - Skończona przestrzeń wartości $Y = \{1, 2, 3, \dots, \bar{Y}\}$
 - Odpowiadająca funkcja gęstości prawdopodobieństwa $f_y(\cdot)$
- Wybieramy N niezależnych próbek z funkcją gęstości: $\{y^1, y^2, y^3, \dots, y^N\}$
- Dla $n \in \{1, 2, \dots, N\}$ niechaj

$$p_i^n := \begin{cases} 1 & \text{jeśli } y^n = i \\ 0 & \text{w pozostałych przypadkach} \end{cases}, \quad i = 1, \dots, \bar{Y}$$

- Niech p_i oznacza średnią z p_i^n na N próbkach

$$p_i := \frac{1}{N} \sum_{n=1}^N p_i^n$$

- Z prawa wielkich liczb

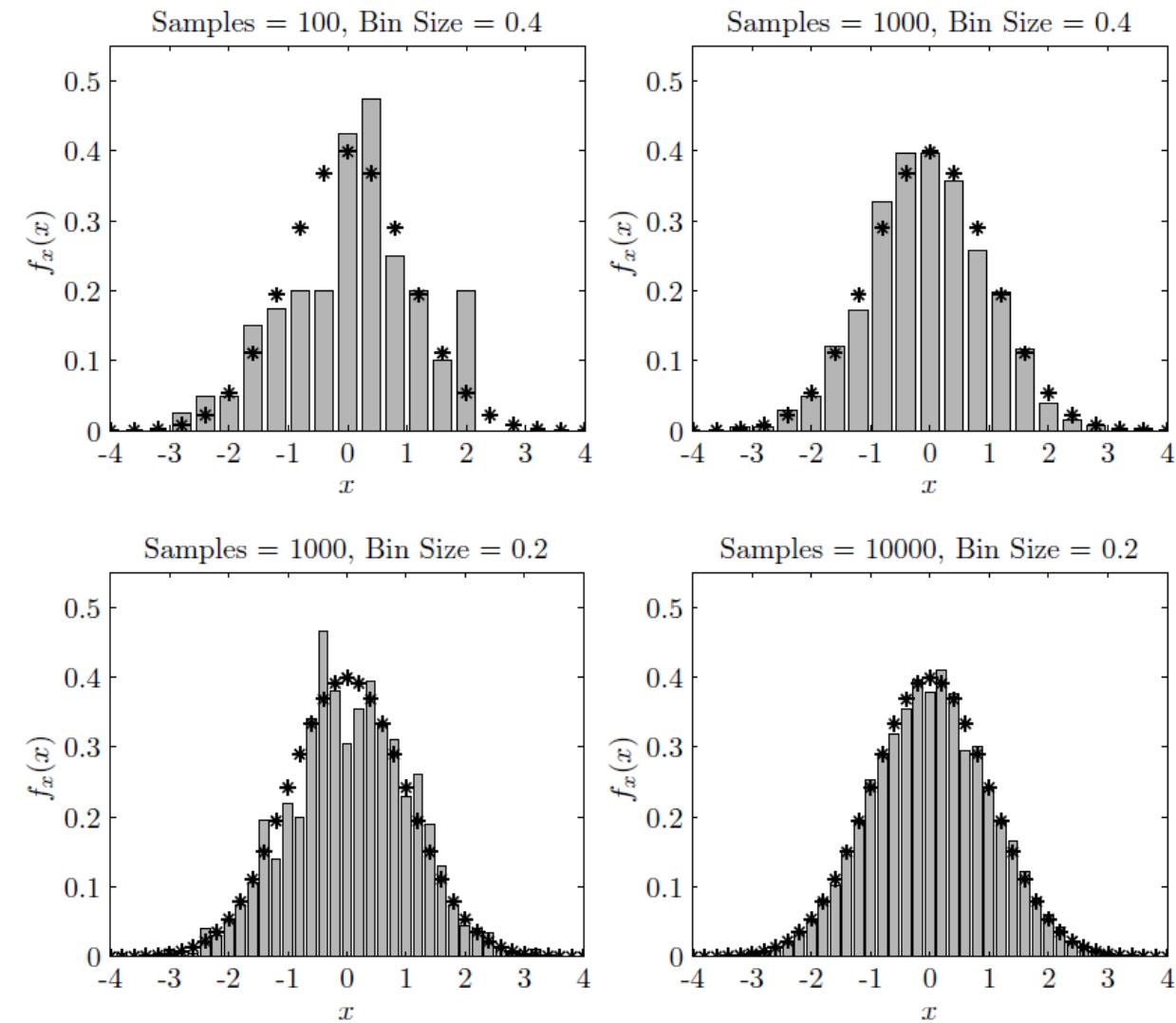
$$p_i \xrightarrow{N \rightarrow \infty} E[p_i^n] = f_y(i)$$

- Wniosek:

funkcja gęstości prawdopodobieństwa zmiennej losowej może być aproksymowana powtarzanym próbkowaniem z funkcji gęstości prawdopodobieństwa

Filtr cząstek – Particle Filter przykład

- Aproxymujemy rozkład normalny o zerowej średniej oraz wariancji powtarzając rysowanie próbek z tej funkcji gęstości.
- Poniższy rysunek pokazuje przykłady dla różnych ilości próbek oraz rozmiarów podziałów na podzbiory (ang. bins) Δx .
- Rzeczywisty kształt pokazany jest gwiazdkami, natomiast histogram pokazuje próbowanie Monte Carlo.



Filtr cząstek – Particle Filter

- Bayesowski estymator stanu (filtr śledzący) wyznacza
 $f(x(k)|z(1:k))$
- Celem filtru cząstek jest aproksymacja
 $f(x(k)|z(1:k))$
- Wyznaczamy dodatkowe zmienne jak dla filtru Kalmana: $x_p(k)$, $x_m(k)$ i $z_m(k)$.
- Rozróżniamy zmienną $z(k)$ oraz wartość $\bar{z}(k)$ jaką ona przyjmuje (aktualny pomiar w chwili k).

$$\left. \begin{array}{l} \text{Init: } x_m(0) := x(0) \\ \text{S1: } x_p(k) := q_{k-1}(x_m(k-1), v(k-1)) \\ \text{S2: } z_m(k) := h_k(x_p(k), w(k)) \\ x_m(k) \text{ defined via its PDF} \\ f_{x_m(k)}(\xi) := f_{x_p(k)|z_m(k)}(\xi|\bar{z}(k)) \quad \forall \xi \end{array} \right\} k = 1, 2, \dots$$

$$f_{x_p(k)}(\xi) = f_{x(k)|z(1:k-1)}(\xi|\bar{z}(1:k-1))$$

$$f_{x_m(k)}(\xi) = f_{x(k)|z(1:k)}(\xi|\bar{z}(1:k)).$$

Filtr cząstek – Particle Filter

predykcja *a priori*

- Zadanie:
mając funkcję gęstości $f_{x_m(k-1)}(\cdot)$ zmiennej $x_m(k - 1)$ budujemy funkcje gęstości $f_{x_p(k)}(\cdot)$ zmiennej $x_p(k)$. Obie będziemy aproksymować metodą próbkowania Monte Carlo.

- Niech:

$$f_{x_m(k-1)}(\xi) \approx \frac{1}{N} \sum_{n=1}^N \delta(\xi - x_m^n(k-1)) \quad \forall \xi$$

- gdzie $\{x_m^n(k - 1)\}$ to jest N próbek aproksymujących funkcje gęstości $x_m(k - 1)$. Są to próbki Monte Carlo $f_{x_m(k-1)}(\cdot)$.

$$f_{x_p(k)}(\xi) \approx \frac{1}{N} \sum_{n=1}^N \delta(\xi - x_p^n(k)) \quad \forall \xi$$

- gdzie $x_p^n(k) := q_{k-1}(x_m^n(k-1), v^n(k-1))$, dla $n=1,2,\dots,N$ i $v^n(k-1)$ jest próbka $f_{v(k-1)}(\cdot)$.

Filtr cząstek – Particle Filter

predykcja *a priori*

- Metoda jest intuicyjna. Propagujemy N cząstek przez dynamikę procesu. Taka symulacja równoległa.
- Założywszy N wystarczająco „duże” i $\{x_m^n(k-1)\}$ jako „*dobrą*” reprezentację $f_{x_m(k-1)}(\cdot)$, to wtedy $\{x_p^n(k)\}$ będzie „*dobrą*” reprezentacją $f_{x_p(k)}(\cdot)$.

Filtr cząstek – Particle Filter

aktualizacja *a posteriori*

- Zadanie:
mając funkcję gęstości $f_{x_p(k)}(\cdot)$ zmiennej $x_p(k) \in X$ (z predykcji *a priori* pomiaru $\bar{z}(k)$) budujemy funkcje gęstości $f_{x_m(k)}(\cdot)$ zmiennej $x_m(k)$. Obie będziemy aproksymować metodą próbkowania Monte Carlo.
- Z reguły Bayesa mamy:

$$f_{x_m(k)}(\xi) := f_{x_p(k)|z_m(k)}(\xi|\bar{z}(k)) = \frac{f_{z_m(k)|x_p(k)}(\bar{z}(k)|\xi) \ f_{x_p(k)}(\xi)}{\sum_{\zeta \in \mathcal{X}} f_{z_m(k)|x_p(k)}(\bar{z}(k)|\zeta) \ f_{x_p(k)}(\zeta)}, \quad \forall \xi.$$

- Aproksymujemy $f_{x_p(k)}(\cdot)$ na podstawie próbkowania Monte Carlo

$$f_{x_p(k)}(\xi) \approx \frac{1}{N} \sum_{n=1}^N \delta(\xi - x_p^n(k)), \quad \forall \xi$$

Filtr cząstek – Particle Filter

aktualizacja *a posteriori*

- Podstawiając $f_{x_p(k)}(\cdot)$ do wzoru na $f_{x_p(k)|z_m(k)}(\xi|\bar{z}(k))$ otrzymujemy

$$f_{x_p(k)|z_m(k)}(\xi|\bar{z}(k)) \approx \sum_{n=1}^N \beta_n \delta(\xi - x_p^n(k)) = \alpha f_{z_m(k)|x_p(k)}(\bar{z}(k)|\xi) \sum_{n=1}^N \delta(\xi - x_p^n(k))$$

gdzie α jest stałą dla wszystkich ξ .

- Otrzymujemy N równań podstawiając odpowiednio $\xi = x_p^n(k)$ dla $n=1,2,\dots,N$ (pamiętajmy, że $\delta(\cdot)$ wynosi zero, oprócz sytuacji gdy argument jest zerowy)

$$\beta_n = \alpha f_{z_m(k)|x_p(k)}(\bar{z}(k)|x_p^n(k)) \quad n = 1, 2, \dots, N$$

- Ponieważ wymagamy $\sum_{n=1}^N \beta_n = 1$, aby $f_{x_p(k)|z_m(k)}(\xi|\bar{z}(k))$ było ważną funkcją gęstości prawdopodobieństwa, otrzymujemy:

$$\alpha = \left(\sum_{n=1}^N f_{z_m(k)|x_p(k)}(\bar{z}(k)|x_p^n(k)) \right)^{-1}$$

Filtr cząstek – Particle Filter

aktualizacja *a posteriori*

- Intuicyjnie traktujemy każdą cząstkę oddziennie. Aktualizacja *a posteriori* używa tych samych cząstek co predykcja *a priori*, tyle że skalowanych wiarygodnością pomiaru.
- Podsumowanie: otrzymujemy reprezentację funkcji gęstości $f_{x_m(k)}(\cdot)$

$$f_{x_m(k)}(\xi) = f_{x_p(k)|z_m(k)}(\xi|\bar{z}(k)) \approx \sum_{n=1}^N \beta_n \delta(\xi - x_p^n(k)), \quad \forall \xi$$

- Chociaż to jeszcze nie koniec. Cząstki muszą być identycznie ważone $1/N$.

Filtr cząstek – Particle Filter

aktualizacja *a posteriori*

- Ponowne próbkowanie
 - Wybieramy N cząstek, z prawdopodobieństwem wybrania n-tej cząstki β_n . Rysujemy N próbek z funkcji gęstości danej jako $\sum_{n=1}^N \beta_n \delta(\xi - x_p^n(k))$. Stosujemy algorytm:
 - Powtórz N razy:
 1. Wybierz losowy numer r równomiernie z przedziału $(0,1)$
 2. Wybierz cząstkę \bar{n} , taką że $\sum_{n=1}^{\bar{n}} \beta_n \geq r$ i $\sum_{n=1}^{\bar{n}-1} \beta_n < r$
 - Otrzymujemy N nowych cząstek $x_m^n(k)$, będących podzbiorem poprzednich. Wszystkie mają taką samą wagę.

$$f_{x_m(k)}(\xi) \approx \frac{1}{N} \sum_{n=1}^N \delta(\xi - x_m^n(k)), \quad \forall \xi$$

- To kończy się praca algorytmu.

Filtr cząstek – Particle Filter

podsumowanie

- **Inicjalizacja:** wybieramy N próbek $\{x_m^n(0)\}$ z $f_{x(0)}(\cdot)$. To są wstępne cząstki.
- **Krok 1: Predykcja *a priori***

Wyznaczamy cząstki *a priori* $\{x_p^n(k)\}$ na podstawie równania stanu dla znanych cząstek $\{x_m^n(k-1)\}$

$$x_p^n(k) = q_{n-1}(x_m^n(k-1), v^n(k-1)) \text{ dla } n=1,2,\dots,N,$$

co wymaga N próbek szumu z rozkładu $f_{v(k-1)}(\cdot)$

- **Krok 2: Aktualizacja *a posteriori***

Skalujemy wszystkie próbki przez wiarygodność pomiaru

$$\beta_n = \alpha f_{z(k)|x(k)}(\bar{z}(k)|x_p^n(k)) \text{ dla } n=1,2,\dots,N,$$

Gdzie α jest stała normalizacji, taką że $\sum_{n=1}^N \beta_n = 1$.

Wykonujemy ponowne próbkowanie w celu otrzymania N nowych cząstek *a posteriori* $x_m^n(k)$.



Filtr cząstek – Particle Filter

problemy

- Wszystkie cząstki zbiegają do jednej → co jest błędnią reprezentacją funkcji gęstości prawdopodobieństwa. Wynika ze skończonej ilości posiadanych danych N (założenie było, że $N \rightarrow \infty$).
- Rozwiążanie jest bardzo kosztowne obliczeniowo. W szczególności konieczność posiadania dużej ilości próbek N .
- Przedstawiony algorytm jest w wersji podstawowej. Istnieje wiele jego modyfikacji, w szczególności uwzględniających zastosowanie wydajniejszych metod numerycznych.



MI

Metody Identyfikacji

wykład #10c

1. *Identyfikacja opóźnienia*

Opóźnienie

- Ogólna definicja – ciągły model nieliniowy z opóźnieniem

$$\frac{d}{dt}x(t) = f(x(t), x(t - \tau_1), \dots, x(t - \tau_k))$$

- Wielokrotne opóźnienia

$$\frac{d}{dt}x(t) = f(x(t), x(t - \tau_1), \dots, x(t - \tau_k), u(t))$$

- W wersji dyskretnej

$$\frac{d}{dt}x(t) = Ax(t) + \sum_{i=1}^k A_i x(t - \tau_i) + u(t)$$

opóźnienia

- Trzy rodzaje opóźnień:
 - Pomiar
 - Wejście
 - Stan
- Opóźnienie stanu jest szczególnie trudne do identyfikacji:
 - Nieskończona wymiarowość
 - Nieróżniczkowalność względem opóźnienia – klasyczne podejścia (LS, ML) nie działają

Identyfikacja opóźnienia

- Założymy najprostszy przypadek.
- Obiekt liniowy, SISO z opóźnieniem
 - $G(s) = G_r(s) \cdot e^{-sTd}$
 - $G(z) = G_r(z) \cdot z^{-d}$
- Zakładamy istnienie zakłóceń w układzie
- Rozważymy sytuacje zarówno otwartej jak i zamkniętej pętli sterowania
- W drugiej części rozważona zostanie sytuacja z modelem nieliniowym

Identyfikacja opóźnienia

zadanie

- Identyfikacja najlepszej możliwej do osiągnięcia aproksymacji opóźnienia
 - np. identyfikacja opóźnienia w interesującym nas paśmie częstotliwościowym
- Identyfikacja rzeczywistej wartości opóźnienia

Właściwości

- Czyste opóźnienie e^{-sTd} jest systemem liniowym
- Czyste opóźnienie e^{-sTd} obejmuje całe pasmo częstotliwości
- Obiekt liniowy ciągły z opóźnieniem jest systemem o nieskończonym rzędzie: nieskończona ilość wartości jest potrzebna do jego opisania w każdej chwili czasu
- Transmitancja czystego opóźnienia nie jest rzeczywista, gdyż ma nieskończoną ilość biegunów
- System w wersji z czasem dyskretnym i stałym okresem próbkowania ma skończony rząd (rozmiar) – opis za pomocą równania różnicowego
- System ze zmiennym okresem próbkowania nie może być opisany równaniami różnicowymi.
- Transmitancja różnicowa czystego opóźnienia jest rzeczywista – skończona ilość biegunów

Klasy metod identyfikacji opóźnienia

1. Metody aproksymacji: aproksymacja relacji pomiędzy sygnałem wejściowym i wyjściowym. Opóźnienie nie jest jasno zdefiniowanym parametrem modelu.
 - a. Identyfikacja w dziedzinie czasu:
 - odpowiedź impulsowa
 - maksimum funkcji korelacji wzajemnej wejście-wyjście
 - Metody *przeparametryzowania* wielomianu licznika
 - b. Metody w dziedzinie częstotliwości
 - Poprzez przesunięcie fazowe
 - c. Metody w dziedzinie funkcji ortogonalnych Laguerra (również i inne funkcje, jak Kautz)
Dwa kroki:
 - i. estymacja modelu
 - ii. estymacja opóźnienia z modelu

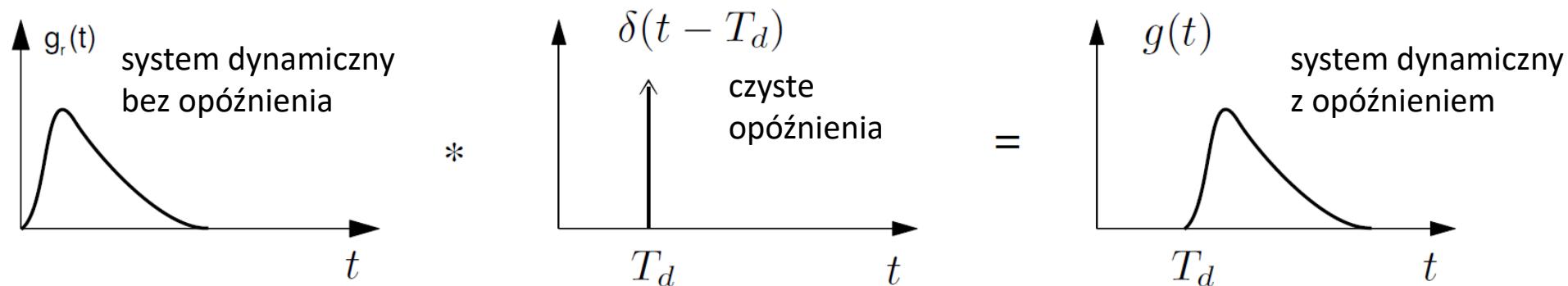
Klasy metod identyfikacji opóźnienia cd.

1. Metody bezpośrednie (*explicit*) – opóźnienie jest jednym z parametrów modelu
 - a. Metody jedno-etapowe: wyznaczamy wszystkie parametry modelu (czyli opóźnienie i pozostałe) w jednym zadaniu identyfikacji
 - Identyfikacja transmitancji z czasem ciągłym, jak np. metoda skoku jednostkowego
 - Identyfikacja transmitancji z czasem dyskretnym, jak np. metoda błędu predykci dla modelu ARX z opóźnieniem
 - b. Metody dwu-etapowe. W jednym kroku opóźnienia, a w drugim inne parametry
 - c. Metody z dyskretyzacją
2. Metody *obszarowe i momentowe*.
 - a. Wykorzystanie zależności pomiędzy pewnymi obszarami pod odpowiedzią skokową a opóźnieniem czy też pomiędzy różnymi momentami tejże odpowiedzi
Dwie części:
 - i. Wyznaczenie odpowiedzi skokowej
 - ii. Estymacja opóźnienia
3. Wykorzystanie statystyk wyższego rzędu

Aproksymacja w dziedzinie czasu

Thresholding methods

- Wyznaczamy opóźnienie jako odległość w dziedzinie czasu do początku niezerowych wartości odpowiedzi impulsowej

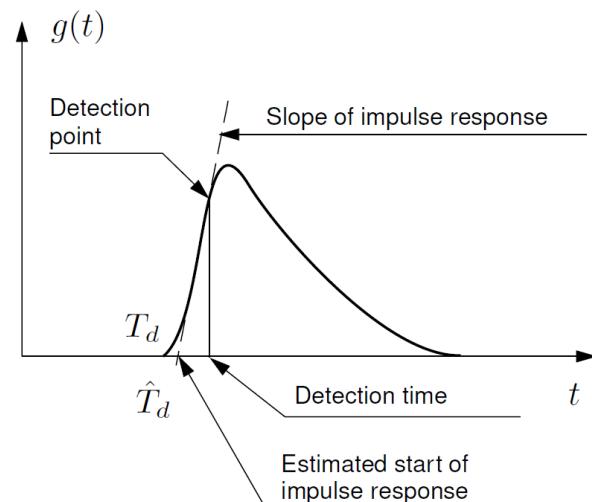


$$g(t) = g_r(t) \cdot \delta(t - T_d)$$

Aproksymacja w dziedzinie czasu

Thresholding methods

- Jeśli mamy dostęp (wyznaczyliśmy) do aproksymacji odpowiedzi impulsowej $\hat{g}(t)$ to możemy wykorzystać:
 1. Metoda separacji
 - a) Separacja opóźnienia $g_d(t)$ od dynamiki $g_r(t)$
 - b) Estymata czasu maximum estymaty odpowiedzi $g_d(t)$
 2. Bezpośrednia detekcja startu odpowiedzi impulsowej
 - a) Detekcja początku zmian odpowiedzi impulsowej (staje się niezerowa)
 - b) Analiza nachylenia odpowiedzi impulsowej



Aproksymacja odpowiedzi impulsowej

- Analiza korelacyjna

$$\hat{g}(\tau) = \frac{\hat{R}_{yu}(\tau)}{\hat{\lambda}}$$

$$\hat{R}_{yu}(\tau) = \frac{1}{N} \sum_{t=r+1}^N y(t)u(t - \tau), \quad \hat{\lambda} = \frac{1}{N} \sum_{t=1}^N u(t)^2$$

- Znalezienie filtra o skończonej odpowiedzi impulsowej (FIR) metoda błędu predykcji

$$\hat{\theta}_N = \underbrace{\left[\frac{1}{N} \sum_t \varphi(t) \varphi(t)^T \right]}_A^{-1} \underbrace{\left[\frac{1}{N} \sum_t \varphi(t) y(t) \right]}_B$$

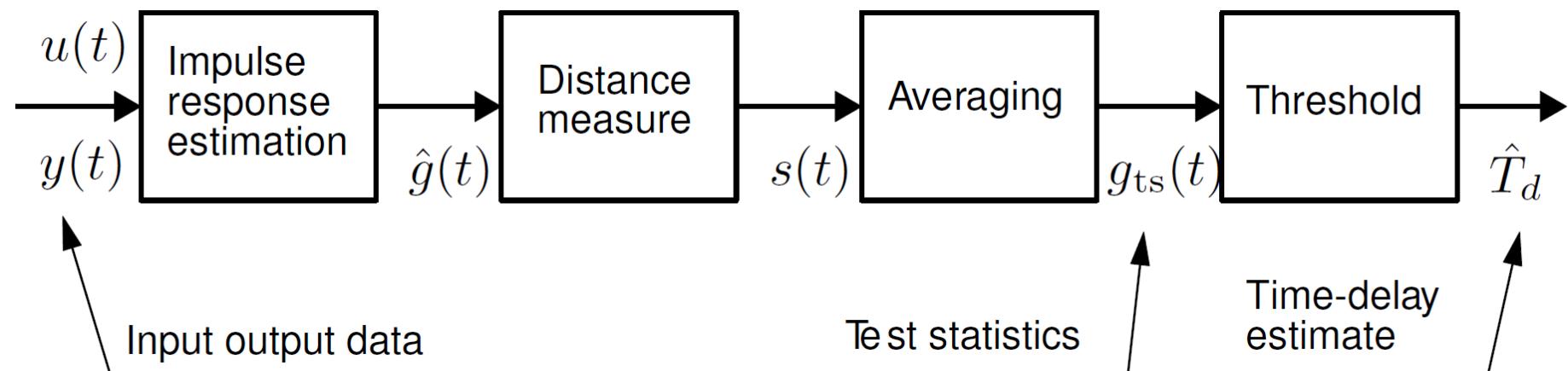
- Jeśli wejście $u(t)$ jest białym szumem, to wtedy macierz $A = \hat{\lambda}I$, $B =$

$$\begin{bmatrix} \hat{R}_{yu}(0) \\ \hat{R}_{yu}(1) \\ \vdots \end{bmatrix}$$

$$\hat{\theta}_N = [\hat{g}_0 \quad \hat{g}_1 \quad \dots]$$



Estymata początku odpowiedzi impulsowej



- Po wyznaczeniu estymaty odpowiedzi impulsowej definiujemy miarę (punkt $s(t)$), podczas gdy odpowiedź zaczyna narastać.
- Redukcja szumów – uśrednianie
- Otrzymujemy statystykę testową $g_{ts}(t)$
- Stosujemy progowanie do wykrycia punktu startu

Uśrednianie

- Odpowiedź skokowa – całkowanie
- Suma kumulacyjna – stosujemy sumowanie kumulacyjne (CUSUM) przed progowaniem

Algorithm 1 CUSUM detector.

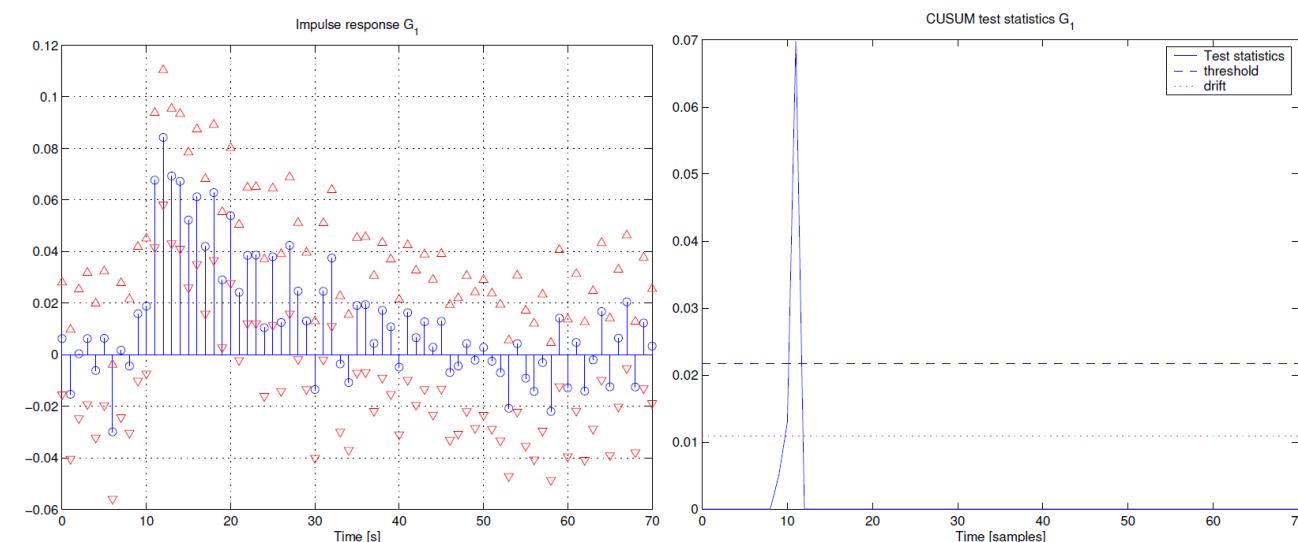
Design parameters: Drift $\nu(t)$ and threshold $h(t)$ that may be time dependent.

Output: Detection time t_a and perhaps the estimate of the change time \hat{k} .

Input: Distance measure $s(t)$.

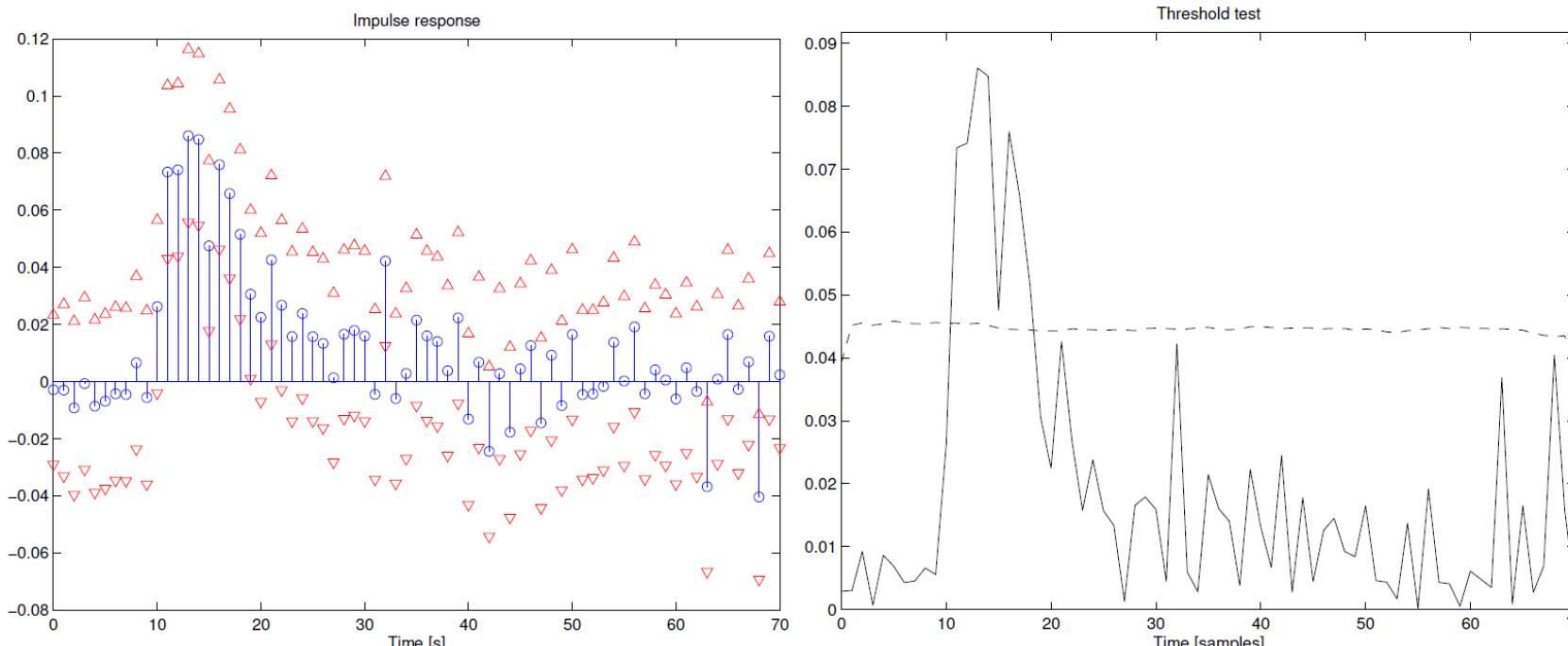
Internal variable: Test statistics $g_{ts}(t)$.

1. $t = 0, g_{ts}(-1) = 0$
2. $g_{ts}(t) = g_{ts}(t - 1) + s(t) - \nu(t)$
3. If $g_{ts}(t) < 0$ then $g_{ts}(t) = 0, \hat{k} = t$
4. If $g_{ts}(t) > h(t) > 0$ then $t_a = t$, goto 6
5. $t = t + 1$, goto 2
6. The detection time is t_a . An estimate of the change time is \hat{k} .



Uśrednianie

- Odpowiedź impulsowa wg Carlemałma
 - Wyznaczamy rekurencyjnie parametry odpowiedzi impulsowej filtrem MNK
 - Poprawiamy estymaty filtrem Kalmana (jeden filtr dla jednego parametru)
 - Wymagania: bały szum na wejściu oraz zakłócenia Gaussowskie
- Metoda bezpośrednia – pomijamy uśrednianie



Progowanie

- Stałe – próg niezmienny a dane niezależne
- Względne progowanie – próg uzależniony jest od niepewności danych
- Metoda progu maksymalnego – (analogia do maksimum funkcji korelacji wzajemnej) – stosujemy gdy mamy czyste opóźnienie (bez dynamiki)
- CFAR (Constant False Alarm Ratio) – popularne w technikach radarowych
- Fault Detection – za pomocą filtra Kalman estymuje po kolejowi wszystkie punkty odpowiedzi skokowej – ten który jest najlepszy daje informacje o opóźnieniu
- Względna metoda Carlemalma – skomplikowana wersja progowania względnego
- Detektor Kurza – detekcja kiedy współczynnika wielomianu licznika estymowanego modelu liniowego przestają być zerowe – wrażliwa na szумy i zakłócenia

Separacja opóźnienia od dynamiki

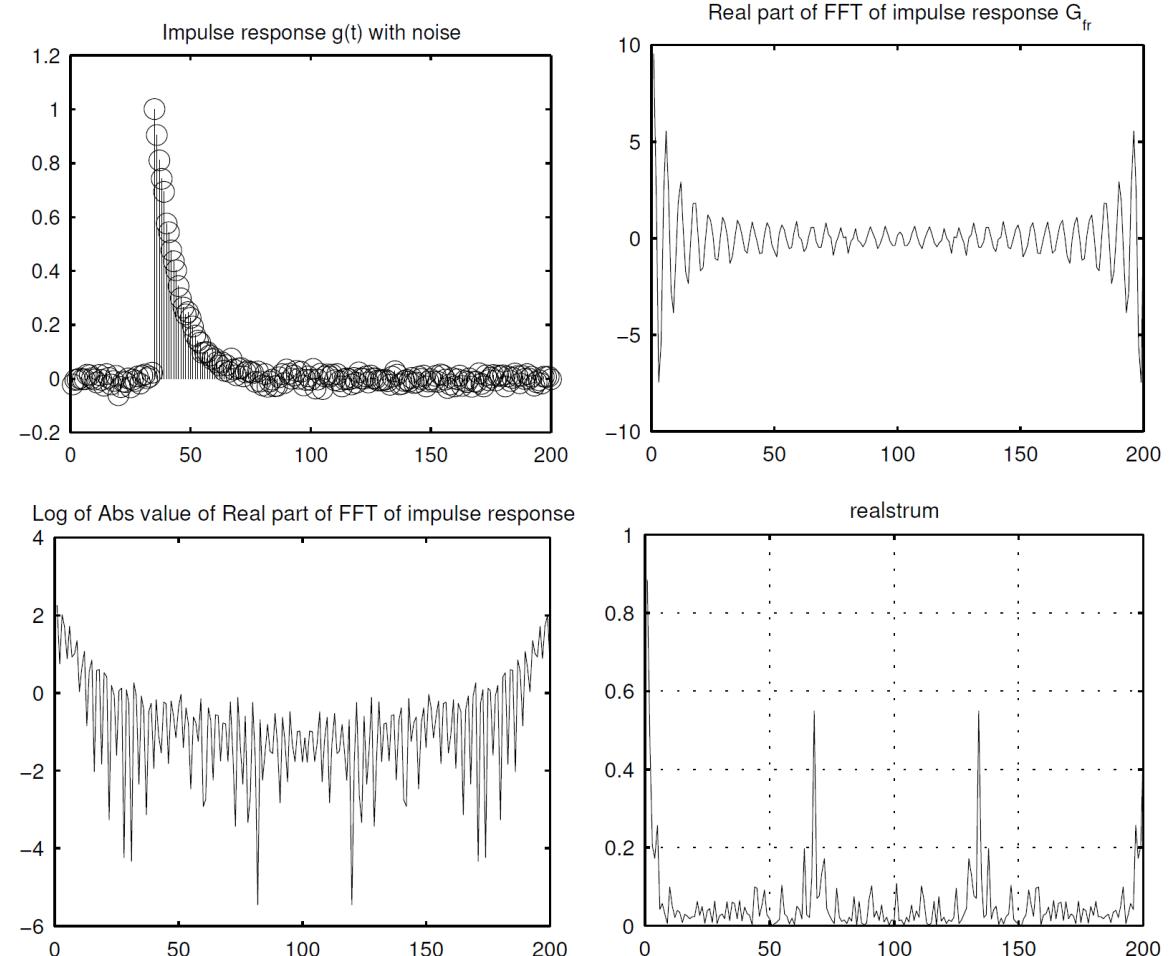
- Liczymy transformację Fouriera odpowiedzi impulsowej

$$G(i\omega) = G_r(i\omega) \cdot e^{-i\omega T_d}$$

- Bierzemy część rzeczywistą

$$G_{fr}(i\omega) = \operatorname{Re}\{G(i\omega)\} = \operatorname{Re}\{G_r(i\omega) \cdot e^{i\omega T_d}\} = A_{fr}(i\omega) \sin(T_d\omega + \varphi(i\omega))$$

- Wyznaczamy logarytm części rzeczywistej transformaty Fouriera
- Liczymy odwrotną transformację Fouriera
- Szukamy pików charakterystyki, które powinny być w punkcie $2T_d$
- Wejście musi być białym szumem, w szczególności nie może oscylować



Metody w dziedzinie częstotliwości – fazowe

- Podstawowa zależność: opóźnienie w dziedzinie czasu $e^{-j\omega T_d}$ odpowiada przesunięciu fazowemu $-\omega T_d$ w dziedzinie częstotliwości.
- Wyznaczamy nachylenie fazy spectrum wzajemnego w dziedzinie częstotliwości
- W wypadku dyskretnego czasu metoda będzie działała do mniej więcej $1/10$ częstotliwości próbkowania
- Estymata charakterystyki częstotliwościowej:

$$\hat{G}(e^{i\omega T}) = \frac{\hat{\Phi}_{yu}(\omega)}{\hat{\Phi}_u(\omega)}$$

- Analizujemy fazę tej charakterystyki w celu wyznaczenia opóźnienia

$$\hat{T}_d = - \left. \frac{d}{d\omega} \arg \hat{G}(e^{i\omega T}) \right|_{\omega=0}$$



Metody w dziedzinie częstotliwości – fazowe

- Zadanie wyznaczenia ch-ki częstotliwościowej było rozważane wcześniej
- Zadanie odpowiada ch-ce korelacyjnej w dziedzinie czasu

$$\Phi_{yu}(\omega) = T_s \sum_{k=-\infty}^{\infty} R_{yu}(k) e^{i\omega k T_s}$$

- Ponieważ opóźnienie jest zależne od pochodnej ch-ki fazowej

$$\frac{d\bar{\varphi}}{d\omega} = \frac{d \arg \bar{G}_{ap}(i\omega)}{d\omega} = \frac{d}{d\omega} \arg e^{-i\omega T_d} = \frac{d}{d\omega} \{-\omega T_d\} = -T_d$$

- Wyznaczamy pochodną ch-ki fazowej

$$T_d \approx -\frac{\varphi(\omega_1 T_s)}{\omega_1 T_s} = -\frac{\arg G_{ap}(e^{i\omega_1 T_s})}{\omega_1 T_s}$$

metoda DAP (Discrete-time Alpass Phase)

Metody w dziedzinie funkcji Laguerra

modele dyskretne

- Postać dyskretna funkcji Laguerra w dziedzinie częstotliwości:

$$L_k(q) = \frac{\sqrt{(1 - \alpha^2)T_s}q^{-1}}{1 - \alpha q^{-1}} \left(\frac{-\alpha + q^{-1}}{1 - \alpha q^{-1}} \right)^{k-1}$$

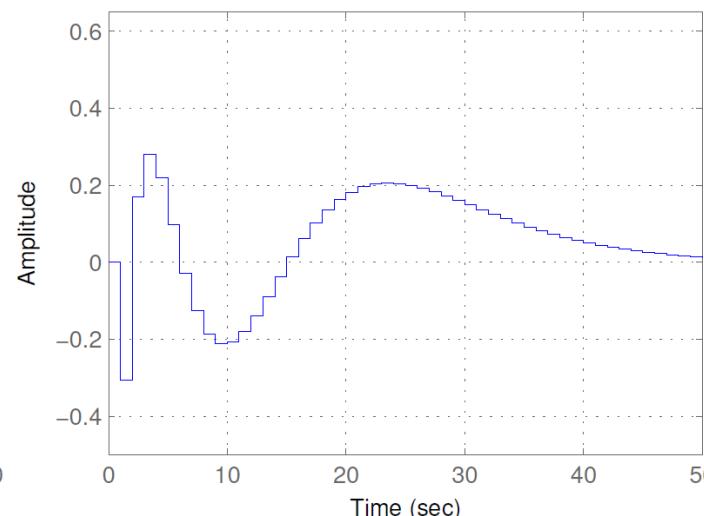
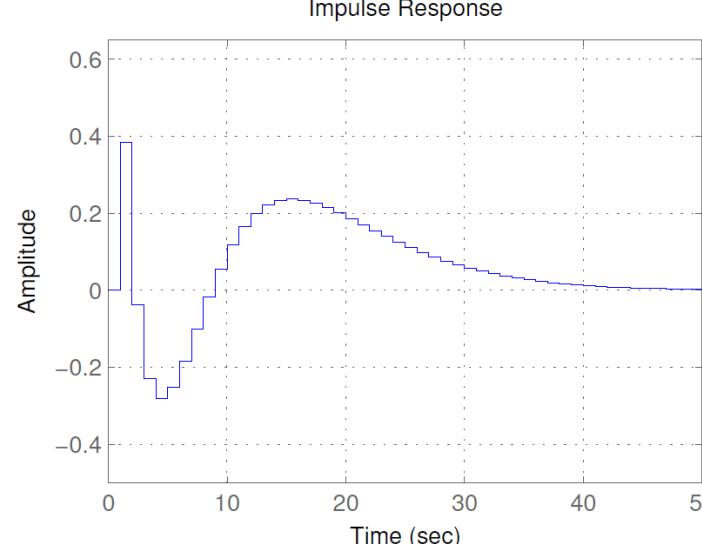
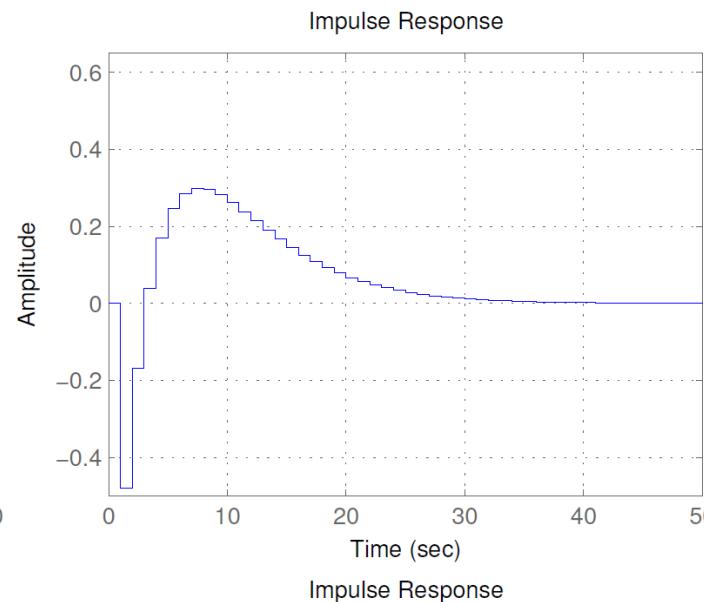
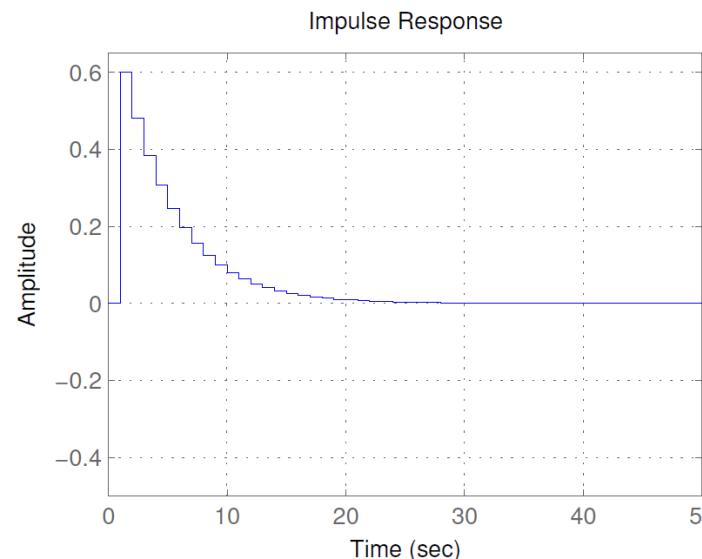
- Po przejściu w dziedzinę czasu otrzymujemy

$$l_k = \mathcal{Z}^{-1} \{ L_k(z) \}$$

Metody w dziedzinie funkcji Laguerra

modele dyskretne

Przykładowe pierwsze cztery funkcje, $\alpha = 0.8$, $T_s = 1$.



$$w(t) = \sum_{k=0}^{\infty} w_k l_k(t)$$

Metody w dziedzinie funkcji Laguerra

modele dyskretne

- Model dyskretny (t, d : dyskretne)

$$y(t) = u(t - d) + n(t)$$

- W dziedzinie funkcji Laguerra otrzymujemy

$$y_k = \sum_{l=0}^k \phi_{k,l} u_l + n_k$$

gdzie y_k, u_l, n_k są współczynnikami Laguerra dla wyjścia, wejścia i szumu.

- Zależność przepisujemy do postaci regresyjnej

$$y_k = \varphi_k^T \Theta + n_k, \quad \varphi_k = [\varphi_{k,1}, \dots, \varphi_{k,N+1}]^T$$

$$\varphi_{k,1} = u_0$$

$$\varphi_{k,l+1} = \frac{(1-\alpha^2)^l}{l!(l-1)!} \sum_{m=0}^{k-l} (-1)^{k+l-m} \alpha^{k-2l-m} \frac{(k-m-1)!}{(k-m-l)!} u_m, \quad k \geq l > 0$$

$$\varphi_{k,l+1} = 0 \quad N \geq l > k$$

$$\Theta = [1, \quad d, \quad \dots, \quad d(d-1) \cdots (d-(N-1))]^T \alpha^d$$

Metody w dziedzinie funkcji Laguerra

modele dyskretne

- Przykładowe rozwiązanie:

$$Y = [y_0, \dots, y_N]^T$$

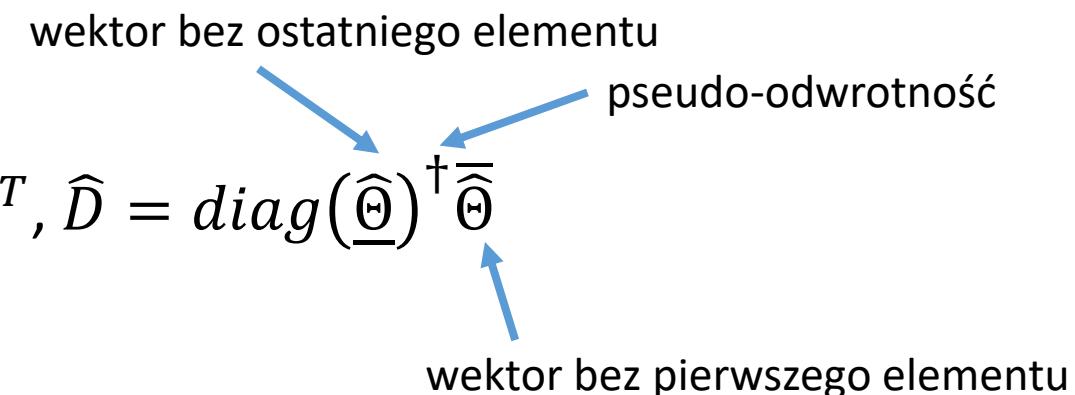
$$\Phi = [\varphi_0, \dots, \varphi_N]^T$$

$$\widehat{\Theta} = \Phi^{-1}Y$$

$$\widehat{d} = (1^T 1)^{-1} 1^T Y_D = \frac{1}{N} 1^T \widehat{D} + \frac{N-1}{2}$$

gdzie:

$$1 = [1, \dots, 1]^T, Y_D = \widehat{D} + [0, 1, \dots, N-1]^T, \widehat{D} = \text{diag}(\widehat{\Theta})^\dagger \widehat{\Theta}$$



Funkcje Laguerra dopasowujemy optymalizacyjnie

Metody bezpośrednie

metody jednokrokowe

- Zakładamy postać modelu z opóźnieniem:
 - jedno-inercyjna: $G(s) = \frac{K}{1+sT} e^{-sL}$
 - dwu-inercyjna: $G(s) = \frac{K}{(1+sT_1)(1+sT_2)} e^{-sL}$
- Dopasowujemy parametry modelu metodą błędu predykcji czy też największej wiarygodności (optymalizacja iteracyjna)
- Problem o wielu ekstremach:
 - Metoda wielostartowa
 - Optymalizacja globalna
- Analogicznie postępujemy w wersji dyskretnej
- Często stosujemy filtrowanie danych

Metody bezpośrednie

metody dwukrokowe

- **Krok 1:** dyskretyzacja
- **Krok 2:** dopasowanie metodami jak dla modeli regresyjnych

Metody powierzchni

Area method

- Zakładamy postać modelu

- $G(s) = \frac{K}{1+sT} e^{-sL}$ lub $G(s) = \frac{K}{(1+sT)^2} e^{-sL}$

1. Liczymy średni czas przebywania T_{ar}

$$T_{ar} = \frac{A_0}{K}$$

gdzie:

K – wzmocnienie statyczne, A_0 - pole pod odpowiedzią skokową

2. Liczymy pole do czasu T_{ar} , czyli

$$A_1 \int_0^{T_{ar}} s(t) dt$$

3. Dla modelu jedno-inercyjnego mamy $T = \frac{e^1 A_1}{K}$ i opóźnienie $L = T_{ar} - T$

4. Dla modelu dwu-inercyjnego mamy $T = \frac{e^2 A_1}{4K}$ i opóźnienie $L = T_{ar} - 2T$

Metody momentów

Moment method

- Znormalizowana odpowiedź impulsowa
gdzie $h(t)$ to odpowiedź impulsowa

$$f(t) = \frac{h(t)}{\int_0^\infty h(t)dt}$$

$$m_n = \int_0^\infty t^n f(t) dt \text{ moment rzędu } n.$$

- Dla modelu pierwszego rzędu mamy

$$K = \int_0^\infty h(t)dt$$

- Średni czas przebywania: $T_{ar} = m_1$,
- stałą czasową otrzymujemy z równania: $T^2 = m_2 T_{ar}^2$, a opóźnienie $L = T_{ar} - T$
- Dla modelu drugiego rzędu mamy: $T^2 = \frac{1}{2}m_2 - \frac{1}{2}T_{ar}^2$, a opóźnienie $L = T_{ar} - 2T$



Zalecenia

Metoda	Wolny obiekt	Szybki obiekt	Pobudzenie PRBS	Pobudzenie skokowe
Progowanie		×	×	
Metody DAP (częstotliwościowe)		×	×	
W dziedzinie Laguerra	×			×
Dopasowanie z czasem ciągłym (metody błędu predykcji)			×	
Dopasowanie z czasem dyskretnym (metody błędu predykcji)		×	×	
Metody momentów i powierzchniowe	×			

Podsumowanie

- W przypadku otwartej pętli: metody błędu predykcji (ciągłe i dyskretne)
- Metody progowania zawodzą – zakłócenia
 - Problemy z estymacją właściwej odpowiedzi impulsowej
 - Wybielanie sygnałów rzadko pomaga
 - Lepiej działają dla obiektów szybkich i sygnałów typu PRBS
- Metody częstotliwościowe (DAP) są jednymi z lepszych
 - Wrażliwe na szумy
- Metody Laguerra zależą od doboru sygnału wejściowego
 - Zależy czy sygnał wejściowy dobrze odpowiada funkcjom Laguerra
 - Wolne obiekty
 - Długie obliczenia

Podsumowanie

- Metody dopasowania
 - Wynik optymalizacji zależy od punktu startowego
 - Wybielanie sygnałów może pomóc
- Metody powierzchni i momentów mają ograniczony zakres stosowania
 - Problemy z estymacją odpowiedzi czy to skokowej, czy też impulsowej
 - Metody momentów raczej nie mają zastosowania
 - Metody powierzchni są trochę lepsze
 - W ogólności dają lepsze wyniki dla wolnych obiektów, w przeciwieństwie do metod progowych

Metody alternatywne

- Podejście genetyczne (optymalizacyjne):
 - Dopasowanie do odpowiedzi skokowej odpowiedniego modelu: dwa przykłady, tzn. jedno- i dwu-inercyjny z opóźnieniem.
 - Identyfikacja wszystkich parametrów modelu, wraz z opóźnieniem
- Modyfikacja zmiennych stanu i optymalizacja – skomplikowane i mało odporne algorytmy, mnóstwo uznaniowych parametrów

Identyfikacja opóźnienia stanu

- Metody aproksymacyjne (optymalizacyjne) rozwijając opóźnienie poprzez *splajny*
- Metody spektralne, aczkolwiek nieroziwiązany do końca problem wymiarowości
- Podejście poprzez wielokrotne opóźnienia i analizę równań stanu z wielokrotnymi opóźnieniami

$$\dot{\hat{x}} = \sum_{j=0}^m \hat{A}_j \hat{x}(t - \hat{\tau}_j) + \hat{B}_j u(t - \hat{\tau}_j) - \alpha \Delta x(t),$$

- wyrazy dotyczące błędnych opóźnień powinny się zerować