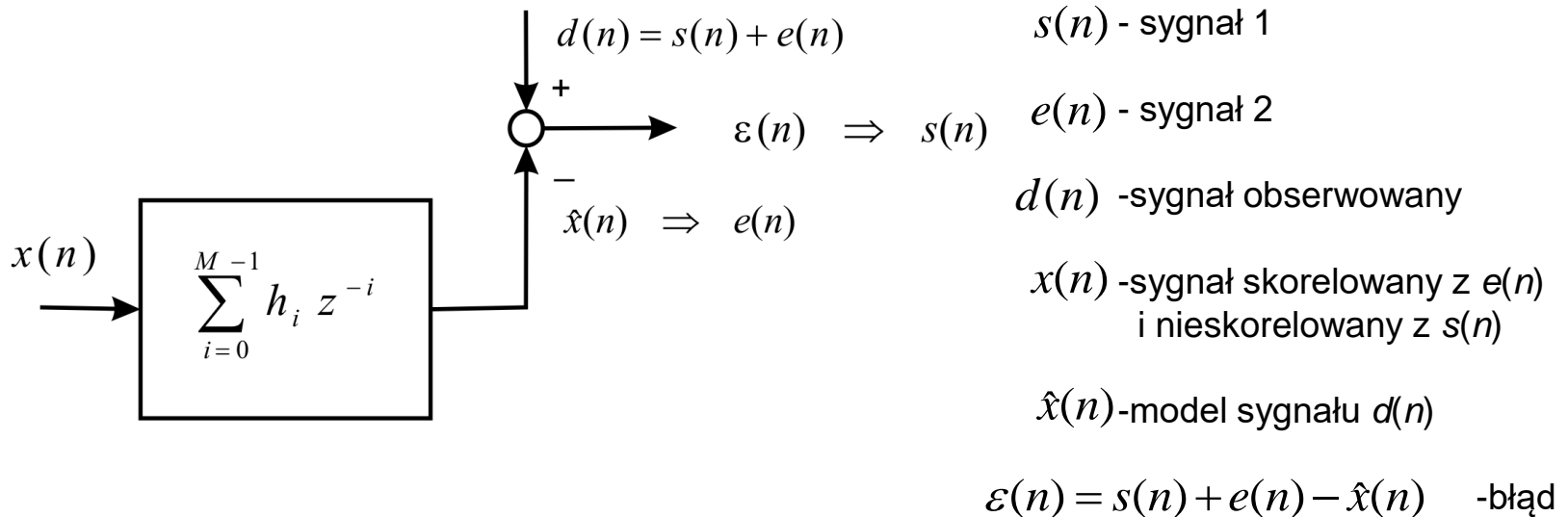


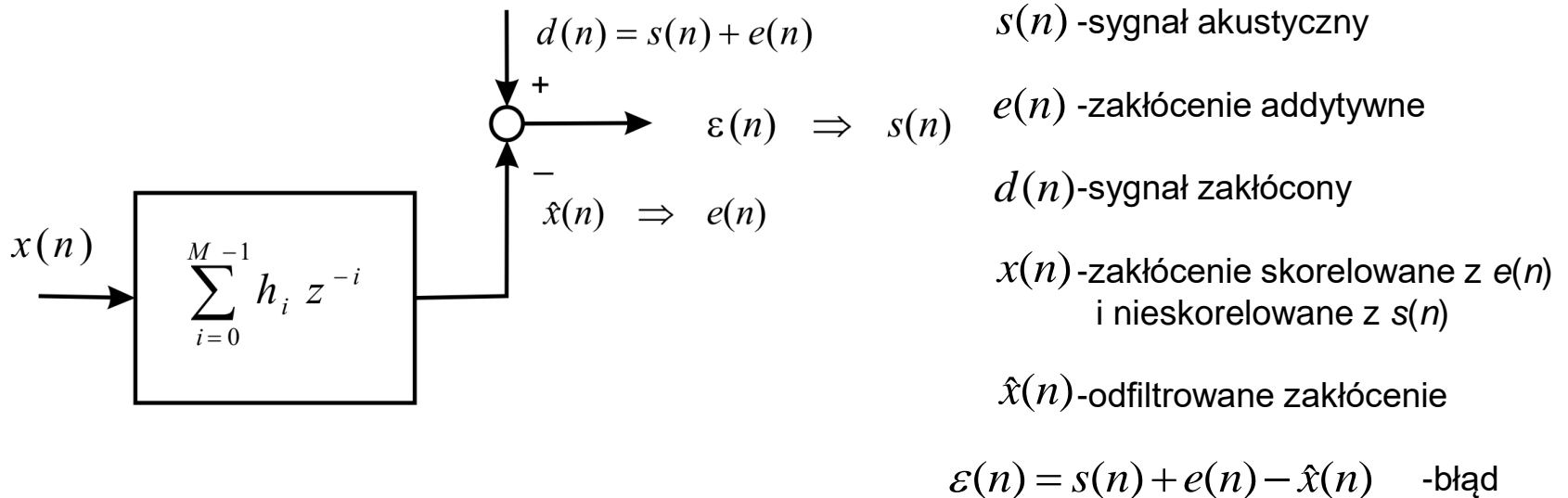
# ZAGADNIENIE W POSTACI OGÓLNEJ



Szukane:  $\bar{h} = [h_0, h_1, \dots, h_{M-1}]^T$  - odpowiedź impulsowa filtru FIR,  $T$  - transpozycja

Kryterium:  $\min E[\varepsilon^2(n)]$

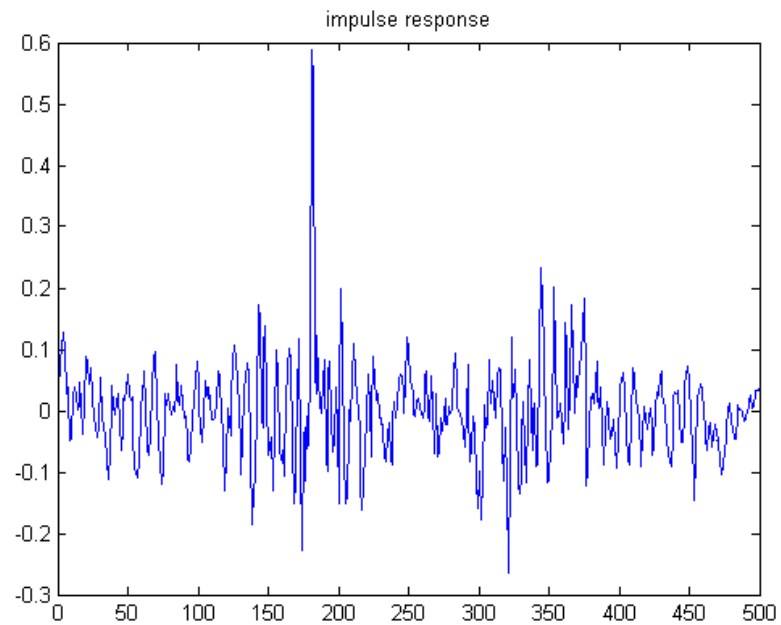
# REDUKCJA ZAKŁÓCEŃ



Szukane:  $\bar{h} = [h_0, h_1, \dots, h_{M-1}]^T$  - odpowiedź impulsowa filtru FIR,  $T$  - transpozycja

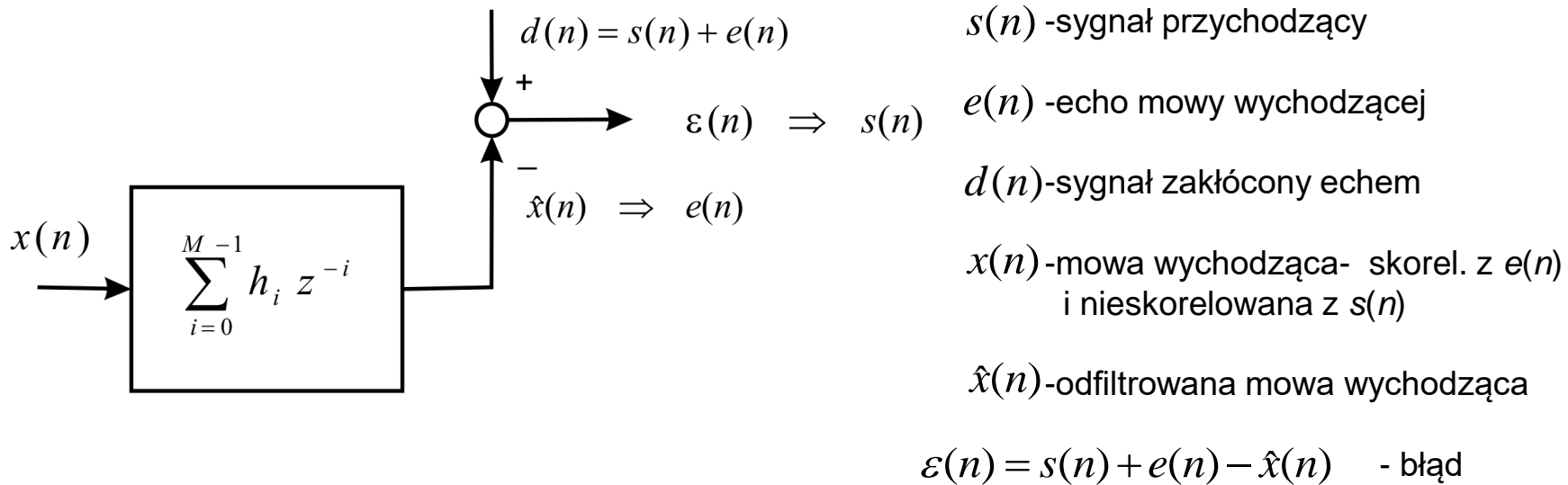
Kryterium:  $\min E[\varepsilon^2(n)]$

# REDUKCJA ZAKŁÓCEŃ - PRZYKŁAD



Przykładowa odpowiedź impulsowa filtru tłumiącego zakłócenie

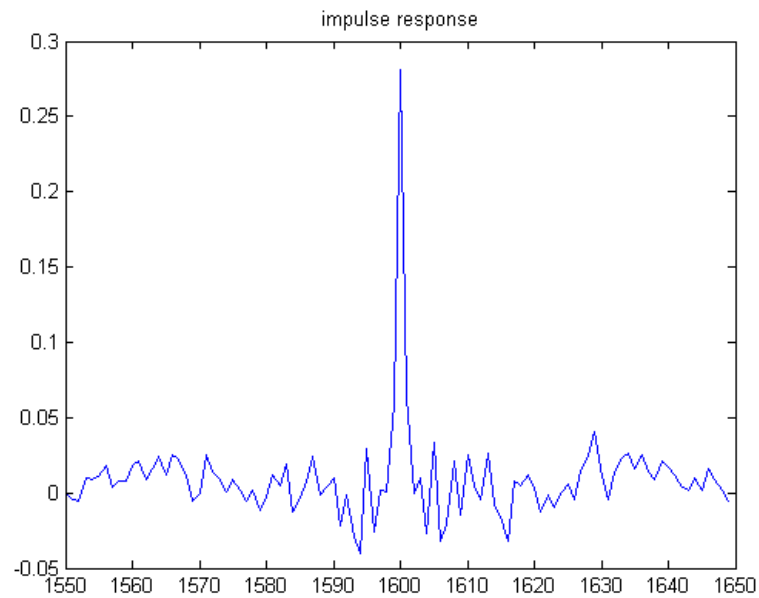
# TŁUMIENIE ECHA



Szukane:  $\bar{h} = [h_0, h_1, \dots, h_{M-1}]^T$  - odpowiedź impulsowa filtru FIR,  $T$  - transpozycja

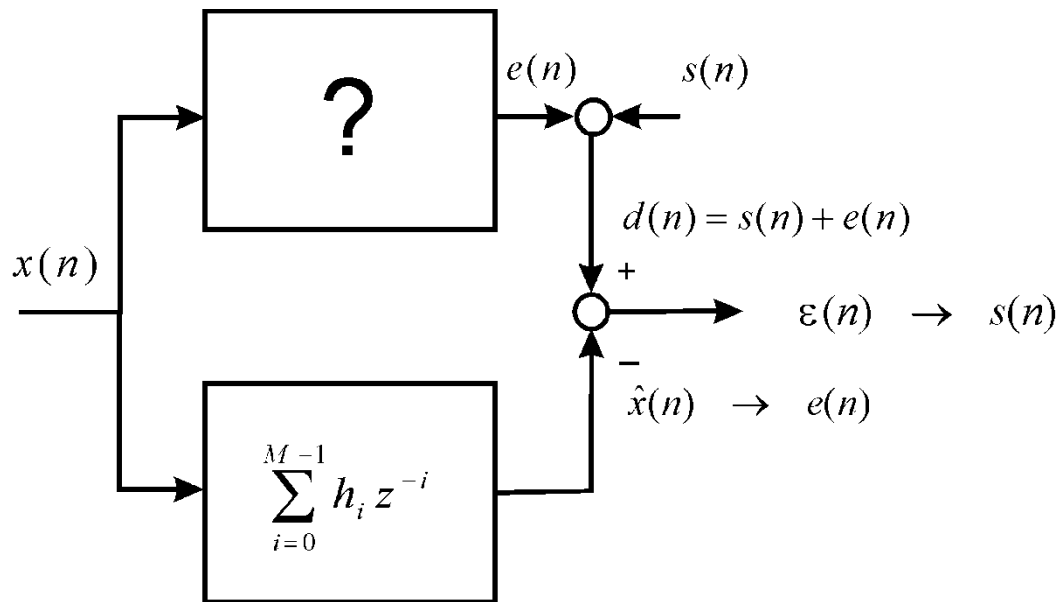
Kryterium:  $\min E[\varepsilon^2(n)]$

# TŁUMIENIE ECHA - PRZYKŁAD



Przykładowa odpowiedź impulsowa filtru tłumiącego echo

# IDENTYFIKACJA OBIEKTU DYNAMICZNEGO



$s(n)$  -szum

$e(n)$  -sygnał wyjściowy obiektu

$d(n)$  -sygnał zakłócony (obserwowany)

$x(n)$  -sygnał pobudzający- skorel. z  $e(n)$   
i nieskorelowany z  $s(n)$

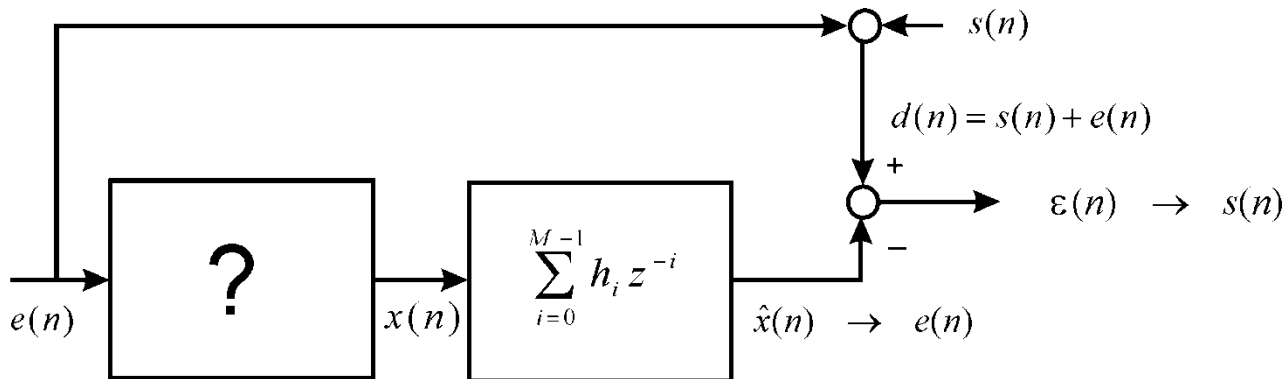
$\hat{x}(n)$  -sygnał wyjściowy modelu

$\varepsilon(n) = s(n) + e(n) - \hat{x}(n)$  - błąd

Szukane:  $\bar{h} = [h_0, h_1, \dots, h_{M-1}]^T$  - odpowiedź impulsowa filtru FIR,  $T$  - transpozycja

Kryterium:  $\min E[\varepsilon^2(n)]$

# KOREKCJA KANAŁU



$s(n)$  - szum  
 $e(n)$  – impulsy (symbole) nadawane  
 $d(n)$  – kopia impulsów nadawanych  
 $x(n)$  - sygnał wyjściowy kanału

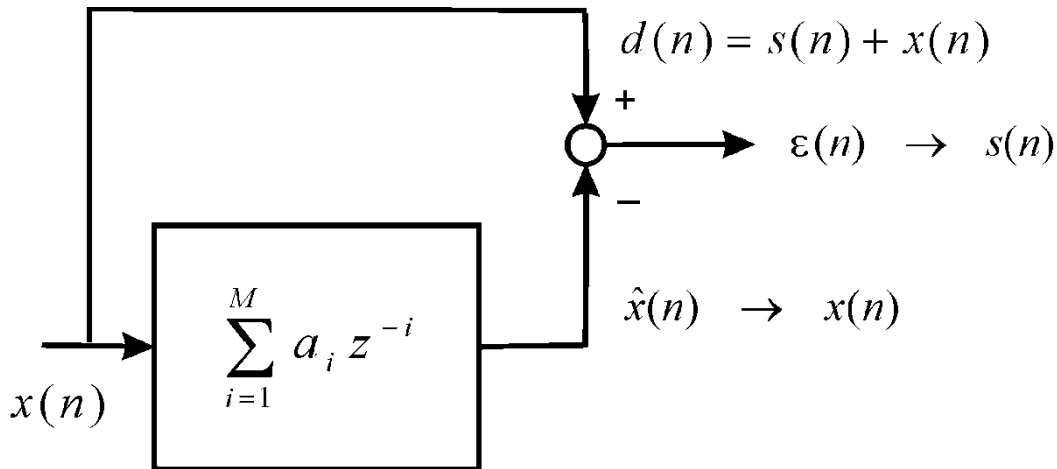
$\hat{x}(n)$  - sygnał wyjściowy korektora

$\varepsilon(n) = s(n) + e(n) - \hat{x}(n)$  - błąd

Szukane:  $\bar{h} = [h_0, h_1, \dots, h_{M-1}]^T$  - odpowiedź impulsowa filtru FIR,  $T$  - transpozycja

Kryterium:  $\min E[\varepsilon^2(n)]$

# PREDYKCJA



$x(n)$  - bieżąca próbka sygnału

$s(n)$  - szum (może być = 0)

$d(n)$  - sygnał obserwowany

$\hat{x}(n)$  - predykcja próbki  $x(n)$

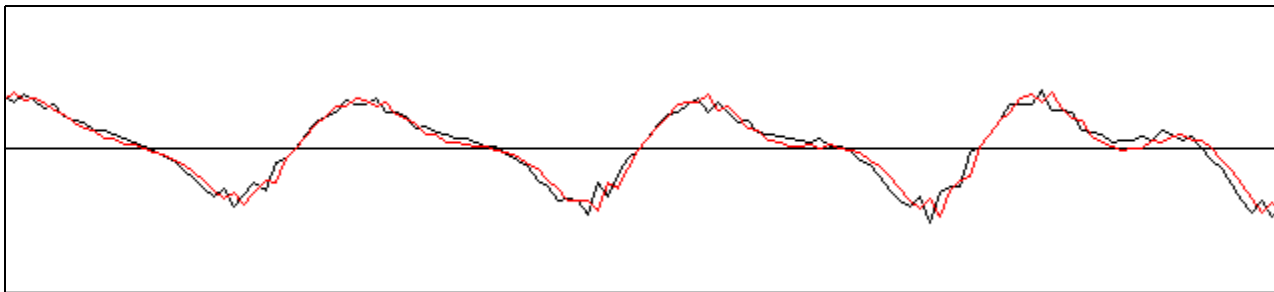
$\varepsilon(n) = s(n) + x(n) - \hat{x}(n)$  - błąd predykcji

Szukane:  $a = [a_1, \dots, a_M]^T$  - współczynniki predykcji,  $T$  - transpozycja

Kryterium:  $\min E[\varepsilon^2(n)]$

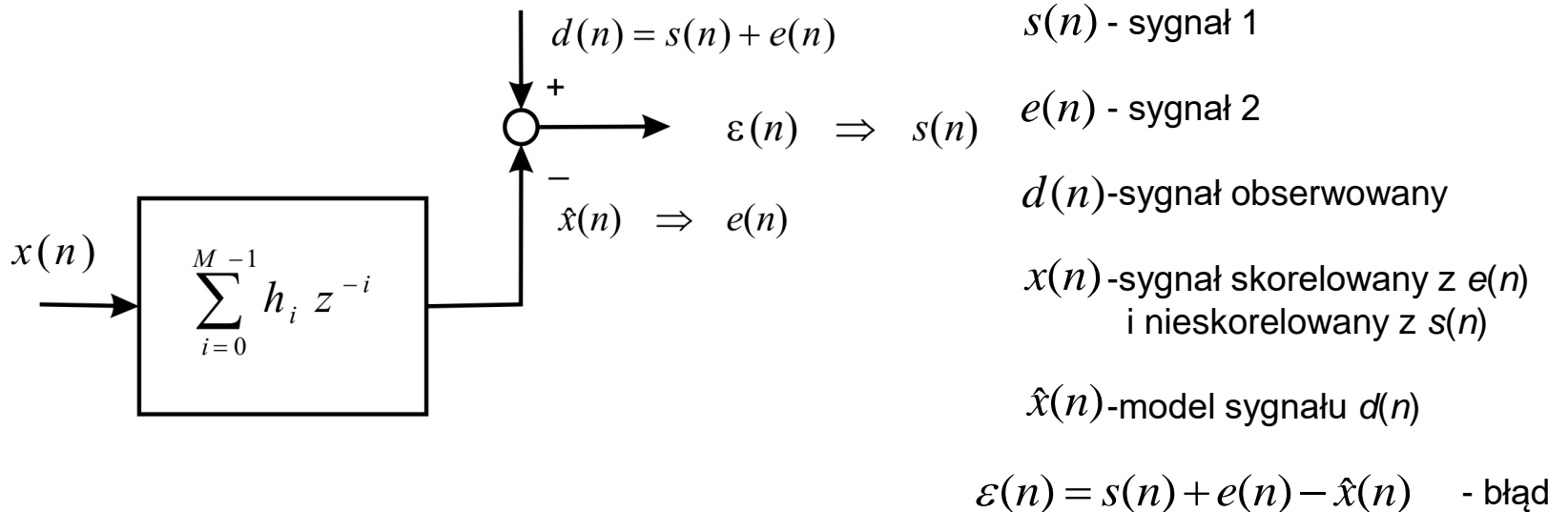


# PREDYKCJA - PRZYKŁAD



Sygnal mowy (czarny) i jego predykcja (czerwony)  
-10 współczynników predykcji

# ZAGADNIENIE W POSTACI OGÓLNEJ



Szukane:  $\bar{h} = [h_0, h_1, \dots, h_{M-1}]^T$  - odpowiedź impulsowa filtru FIR,  $T$  - transpozycja

Kryterium:  $\min E[\varepsilon^2(n)]$

## Rozwiązanie (filtr Wienera)

$$\hat{x}(n) = \sum_{i=0}^{M-1} h_i x(n-i) = \bar{h}^T \bar{x}(n) \quad \text{gdzie} \quad \bar{x}(n) = \begin{pmatrix} x(n) \\ x(n-1) \\ \vdots \\ x(n-M+1) \end{pmatrix}$$

$$\varepsilon(n) = d(n) - \hat{x}(n) = d(n) - \bar{h}^T \bar{x}(n)$$

$$\varepsilon^2(n) = [d(n) - \bar{h}^T \bar{x}(n)]^2 = d^2(n) - 2d(n) \bar{h}^T \bar{x}(n) + \bar{h}^T \bar{x}(n) \bar{x}^T(n) \bar{h}$$

$$E[\varepsilon^2(n)] = E[d^2(n)] - 2\bar{h}^T E[d(n)\bar{x}(n)] + \bar{h}^T E[\bar{x}(n)\bar{x}^T(n)]\bar{h}$$

# Rozwiązanie (filtr Wienera)

$$E[\varepsilon^2(n)] = E[d^2(n)] - 2\bar{h}^T E[d(n)\bar{x}(n)] + \bar{h}^T E[\bar{x}(n)\bar{x}^T(n)]\bar{h}$$

Podstawiamy  $p = E[d(n)x(n)] = \begin{pmatrix} E[d(n)x(n)] \\ E[d(n)x(n-1)] \\ \vdots \\ E[d(n)x(n-M+1)] \end{pmatrix}$

$$\bar{C} = E(x(n)\bar{x}^T(n)) = \{c_{ij}\}_{i,j=0,1,\dots,M-1} = \begin{pmatrix} R_0 & \cdots & R_{M-1} \\ \vdots & \ddots & \vdots \\ R_{M-1} & \cdots & R_0 \end{pmatrix}$$

gdzie  $c_{ij} = E[x(n-i)x(n-j)] = R_{i-j}$  - współczynnik autokorelacji sygnału  $x(n)$

## Rozwiązanie (filtr Wienera)

$$\min E[\varepsilon^2(n)] = \min \{ E[d^2(n)] - 2\bar{h}^T \bar{p} + \bar{h}^T \bar{C} \bar{h} \}$$

Rozwiązanie: 
$$\frac{\partial}{\partial \bar{h}} E[\varepsilon^2(n)] = -2 \frac{\partial}{\partial \bar{h}} [\bar{h}^T \bar{p}] + \frac{\partial}{\partial \bar{h}} [\bar{h}^T \bar{C} \bar{h}] = \bar{0}$$

$$-2\bar{p} + 2\bar{C}\bar{h} = \bar{0} \quad \rightarrow \quad \boxed{\bar{C}\bar{h} = \bar{p}}$$

Dobre wyniki dla sygnału  $s(n)$  i  $x(n)$  o charakterze stacjonarnym

# Adaptacja filtru

1. **Blokowa** – np. dla mowy co 10-30 ms oblicza się nową odpowiedź impulsową

a) obliczenie wsp. autokorelacji sygnału  $x$  i wsp. korelacji sygn.  $d$  i  $x$

$$R_i = \sum_n x(n)x(n-i), \quad p_i = \sum_n d(n)x(n-i) \quad i = 0, 1, \dots, M-1$$

b) obliczenie współczynników filtru – rozwiązanie układu równań liniowych

$$\overline{C} \overline{h} = \overline{p}$$

c) filtracja  $x * \overline{h}$

# Adaptacja filtru

## 2. Sekwencyjna

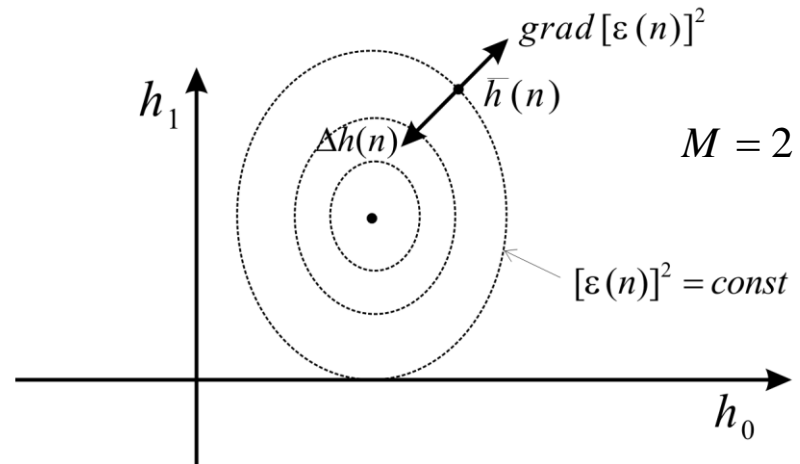
- Co próbkę oblicza się nowe współczynniki filtru, dodając poprawkę do estymaty uzyskanej dla poprzedniej próbki.

$$\bar{h}(n+1) = \bar{h}(n) + \Delta\bar{h}(n) \quad \Delta\bar{h}(n) \quad \text{- poprawka w chwili } n$$

- Poprawki winny powodować zmniejszenie mocy błędu  $[\varepsilon(n)]^2$   
- kryterium dla metod **LMS** (least mean square)

- Metoda gradientowa  
(SG – stochastycznego gradientu)  
- należy do metod LMS

$$\Delta\bar{h}(n) \propto -\text{grad}[\varepsilon(n)]^2$$



# Adaptacja sekwencyjna

Moc chwilowa błędu:  $[\varepsilon(n)]^2 = [d(n) - \bar{h}^T(n)x(n)]^2$

jest funkcją wektora współczynników filtra  $\bar{h}(n)$

Gradient (kierunek wzrostu mocy błędu):

$$\text{grad}[\varepsilon(n)]^2 = \frac{\partial}{\partial \bar{h}} [d(n) - \bar{h}^T(n)x(n)]^2 = -2\varepsilon(n)x(n)$$

Poprawka (w kierunku przeciwnym do gradientu):  $\Delta \bar{h}(n) = -\frac{\beta}{2} \text{grad}[\varepsilon(n)]^2 = \beta \varepsilon(n)x(n)$

**Metoda stochastycznego gradientu:**  $\bar{h}(n+1) = \bar{h}(n) + \Delta \bar{h}(n) = \bar{h}(n) + \beta \varepsilon(n)x(n)$

$\beta$  – szybkość adaptacji



# Metody LS (najmniejszej sumy kwadratów)

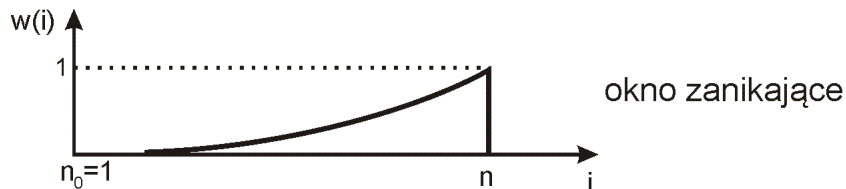
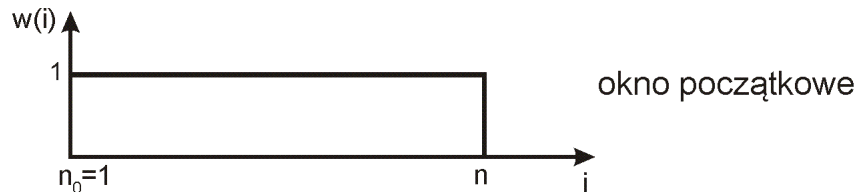
W każdej chwili  $n$  obliczamy nowe współczynniki filtru  $h_0(n), \dots, h_{M-1}(n)$

$$\hat{x}(n) = \sum_{i=0}^{M-1} h_i(n) x(n-i) = \bar{h}^T(n) \bar{x}(n)$$

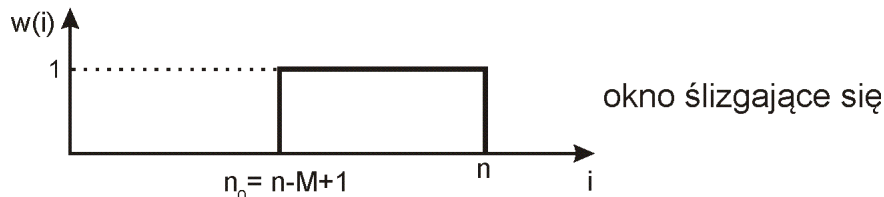
$$\varepsilon(n) = d(n) - \hat{x}(n) = d(n) - \bar{h}^T(n) \bar{x}(n)$$

Kryterium:  $\min_{\bar{h}(n)} E(n) \quad E(n) = \sum_{i=1}^n \varepsilon^2(i)$

(okno początkowe, rozszerzające się)



$$E(n) = \sum_{i=1}^n \lambda^{n-i} \varepsilon^2(i)$$



$$E(n) = \sum_{i=n_0}^n \varepsilon^2(i)$$

# Metody LS (najmniejszej sumy kwadratów)

$$\begin{aligned} E(n) &= \sum_{i=1}^n \varepsilon^2(i) = \sum_{i=1}^n [d(i) - \bar{h}^T(n) \bar{x}(i)]^2 = \\ &= \sum_{i=1}^n d^2(i) - 2 \sum_{i=1}^n d(i) \bar{x}^T(i) \bar{h}(n) + \bar{h}^T(n) \sum_{i=1}^n \bar{x}(i) \bar{x}^T(i) \bar{h}(n) = \\ &= \sum_{i=1}^n d^2(i) - 2 \bar{p}^T(n) \bar{h}(n) + \bar{h}^T(n) \bar{C}(n) \bar{h}(n) \end{aligned}$$

gdzie

$$\bar{p}(n) = \sum_{i=1}^n d(i) \bar{x}(i) \quad \bar{C}(n) = \sum_{i=1}^n \bar{x}(i) \bar{x}^T(i)$$

rozwiązanie

$$\bar{C}(n) \bar{h}(n) = \bar{p}(n)$$

# Wariant sekwencyjny RLS (recursive least squares)

Znamy rozwiązanie  $\bar{C}(n-1)\bar{h}(n-1) = \bar{p}(n-1)$

szukamy  $\bar{h}(n)$ ,  $\bar{C}(n)\bar{h}(n) = \bar{p}(n)$

Okno zanikające (dla  $\lambda=1$  przechodzi w początkowe)

$$E(n) = \sum_{i=1}^n \lambda^{n-i} \varepsilon^2(i) = \sum_{i=1}^n \lambda^{n-i} [d(i) - \bar{h}^T(n) \bar{x}(i)]^2 =$$

$$= \sum_{i=1}^n \lambda^{n-i} d^2(i) - 2 \underbrace{\sum_{i=1}^n \lambda^{n-i} d(i) \bar{x}^T(i)}_{\bar{p}^T(n) \bar{h}(n)} \bar{h}(n) + \bar{h}^T(n) \underbrace{\sum_{i=1}^n \lambda^{n-i} \bar{x}(i) \bar{x}^T(i)}_{\bar{C}(n)} \bar{h}(n) =$$

$$= \sum_{i=1}^n \lambda^{n-i} d^2(i) - 2\bar{p}^T(n)\bar{h}(n) + \bar{h}^T(n)\bar{C}(n)\bar{h}(n)$$

gdzie  $\bar{p}(n) = \sum_{i=1}^n \lambda^{n-i} d(i) \bar{x}(i) = \sum_{i=1}^{n-1} \lambda^{n-i} d(i) \bar{x}(i) + d(n)\bar{x}(n) =$

$$= \lambda \sum_{i=1}^{n-1} \lambda^{n-1-i} d(i) \bar{x}(i) + d(n)\bar{x}(n) = \lambda \bar{p}(n-1) + d(n)\bar{x}(n)$$

podobnie  $\bar{C}(n) = \lambda \bar{C}(n-1) + \bar{x}(n) \bar{x}^T(n)$

# Wariant sekwencyjny RLS (recursive least squares)

Aby uniknąć odwracania macierzy korzystamy ze wzoru:

Jeśli  $A, B$  – macierze kwadratowe,  $x$  – wektor, oraz  $A = B^{-1} + xx^T$

$$\text{to } A^{-1} = B - \frac{B x x^T B}{1 + x^T B x}$$

Podstawiamy:  $A = \bar{C}(n)$ ,  $B^{-1} = \lambda \bar{C}(n-1)$ ,  $x = \bar{x}(n)$

$$\text{Stąd } \bar{C}^{-1}(n) = \frac{1}{\lambda} \bar{C}^{-1}(n-1) - \frac{1}{\lambda} g(n) \bar{x}^T(n) \bar{C}^{-1}(n-1), \quad g(n) = \frac{\bar{C}^{-1}(n-1) \bar{x}(n)}{\lambda(1 + \frac{1}{\lambda} \bar{x}^T(n) \bar{C}^{-1}(n-1) \bar{x}(n))}$$

$$\begin{aligned} \text{Z drugiej strony } \bar{C}^{-1}(n) \bar{x}(n) &= \frac{1}{\lambda} \bar{C}^{-1}(n-1) \bar{x}(n) - \frac{1}{\lambda} g(n) \bar{x}^T(n) \bar{C}^{-1}(n-1) \bar{x}(n) = \\ &= g(n) \left(1 + \frac{1}{\lambda} \bar{x}^T(n) \bar{C}^{-1}(n-1) \bar{x}(n)\right) - \frac{1}{\lambda} g(n) \bar{x}^T(n) \bar{C}^{-1}(n-1) \bar{x}(n) = g(n) \end{aligned}$$

Ostatecznie

$$\begin{aligned} \bar{h}(n) = \bar{C}^{-1}(n) \bar{p}(n) &= \bar{C}^{-1}(n) [\lambda \bar{p}(n-1) + d(n) \bar{x}(n)] = \lambda \bar{C}^{-1}(n) \bar{p}(n-1) + d(n) \bar{C}^{-1}(n) \bar{x}(n) = \\ &= \underbrace{\bar{C}^{-1}(n-1) \bar{p}(n-1)}_{\bar{h}(n-1)} - \underbrace{g(n) \bar{x}^T(n) \bar{C}^{-1}(n-1) \bar{p}(n-1)}_{\bar{h}(n-1)} + d(n) g(n) = \bar{h}(n-1) + g(n) [d(n) - \bar{x}^T(n) \bar{h}(n-1)] \end{aligned}$$

# Wariant sekwencyjny RLS (recursive least squares)

$$\bar{h}(n) = \bar{h}(n-1) + g(n)[d(n) - \bar{x}^T(n)\bar{h}(n-1)]$$

Błąd a priori  $\varepsilon^{pr}(n) = d(n) - \bar{x}^T(n)\bar{h}(n-1)$

Błąd a posteriori  $\varepsilon^{po}(n) = d(n) - \bar{x}^T(n)\bar{h}(n)$

Współczynnik „zapominania”:  $\lambda$  (wpływa na stabilność numeryczną algorytmu RLS)

Start algorytmu:  $n = 1, \quad \bar{C}^{-1}(0) = \delta I, \quad \bar{h}(0) = \bar{0}$

Krok  $n$ : obliczamy  $g(n), \quad \bar{h}(n), \quad \bar{C}^{-1}(n)$

# „Rzadkie” modele

Projektując filtr  $H(z)$  można wykorzystać naszą wiedzę o problemie.

Np. w zagadnieniu tłumienia echa można się spodziewać odpowiedzi impulsowej:

$$\bar{h} = [0, 0, \dots, 0, h_m, h_{m+1}, \dots, h_{M-1}]^T$$

gdzie  $m$  – opóźnienie echa

Długość odpowiedzi impulsowej ( $M$ ) musi być większa niż opóźnienie echa

Podobnie w zagadnieniu identyfikacji odpowiedzi impulsowej kanału wielodrożnego można założyć liczbę dróg ( $K$ ) i szukać odpowiedzi impulsowej:

$$\bar{h} = [0, \dots, 0, h_{j(1)}, 0, \dots, 0, h_{j(2)}, 0, \dots, 0, h_{j(K)}]^T$$

Tworzymy w ten sposób tzw. „rzadkie” modele sygnałów lub układów

# „Rzadkie” modele – c.d.

Przepiszmy równanie błędu dla kolejnych chwil czasowych:

$$\varepsilon(n) = d(n) - \bar{x}^T(n) \bar{h}$$

$$\varepsilon(n-1) = d(n-1) - \bar{x}^T(n-1) \bar{h}$$

.....

$$\varepsilon(1) = d(1) - \bar{x}^T(1) \bar{h}$$



$$\bar{\varepsilon} = \bar{d} - \bar{X} \bar{h}$$

↑  
macierz n x M

W metodach LS minimalizujemy energię błędu:

$$\min_{\bar{h}} \|\bar{\varepsilon}\|^2 = \min_{\bar{h}} \|\bar{d} - \bar{X} \bar{h}\|^2$$

## „Rzadkie” modele – c.d.

$$\min_{\bar{h}} \|\bar{\varepsilon}\|^2 = \min_{\bar{h}} \|\bar{d} - \bar{X}\bar{h}\|^2$$

$$\begin{aligned} \|\bar{d} - \bar{X}\bar{h}\|^2 &= (\bar{d} - \bar{X}\bar{h})^t (\bar{d} - \bar{X}\bar{h}) = \bar{d}^t \bar{d} - 2\bar{d}^t \bar{X}\bar{h} + \bar{h}^t \bar{X}^t \bar{X}\bar{h} = \\ &= \bar{d}^t \bar{d} - 2\bar{p}^t \bar{h} + \bar{h}^t \bar{C}\bar{h} \end{aligned}$$

gdzie  $\bar{p} = \bar{X}^T \bar{d}$ ,  $\bar{C} = \bar{X}^T \bar{X}$

Przyrównując pochodne do zera  $\frac{\partial}{\partial \bar{h}} \|\bar{d} - \bar{X}\bar{h}\|^2 = -2\bar{p} + 2\bar{C}\bar{h} = \bar{0}$

otrzymuje się układ równań liniowych  $\bar{C}\bar{h} = \bar{p}$

Rozwiązując ten układ równań liniowych, otrzymamy najczęściej odpowiedź impulsową, która nie jest „rzadka” (nie zawiera ciągów zer)

→ „rzadkiego” rozwiązania trzeba szukać bezpośrednio, dążąc z  $\bar{X}\bar{h}$  do  $\bar{d}$



# „Rzadkie” modele – matching pursuit (MP)

$$\min_{\bar{h}} \|\bar{\varepsilon}\|^2 = \min_{\bar{h}} \|\bar{d} - \bar{X} \bar{h}\|^2 \quad \bar{X} = [X_1, X_2, \dots, X_M]$$

„Rzadki” model wektora docelowego  $\bar{d}$  :

$$\bar{d}^* = \sum_{i=1}^K h_{j(i)} X_{j(i)}$$

jest kombinacją liniową  $K < M$  kolumn macierzy  $\bar{X}$

## Matching pursuit:

Kolumny wybieramy kolejno,  
stopniowo zmniejszając błąd

$$\bar{\varepsilon} = \bar{d} - \bar{d}^*$$

# „Rzadkie” modele – matching pursuit (MP)

$$\min_{\bar{h}} \|\bar{\varepsilon}\|^2 = \min_{\bar{h}} \|\bar{d} - \bar{X}\bar{h}\|^2 \quad \bar{X} = [X_1, X_2, \dots, X_M]$$

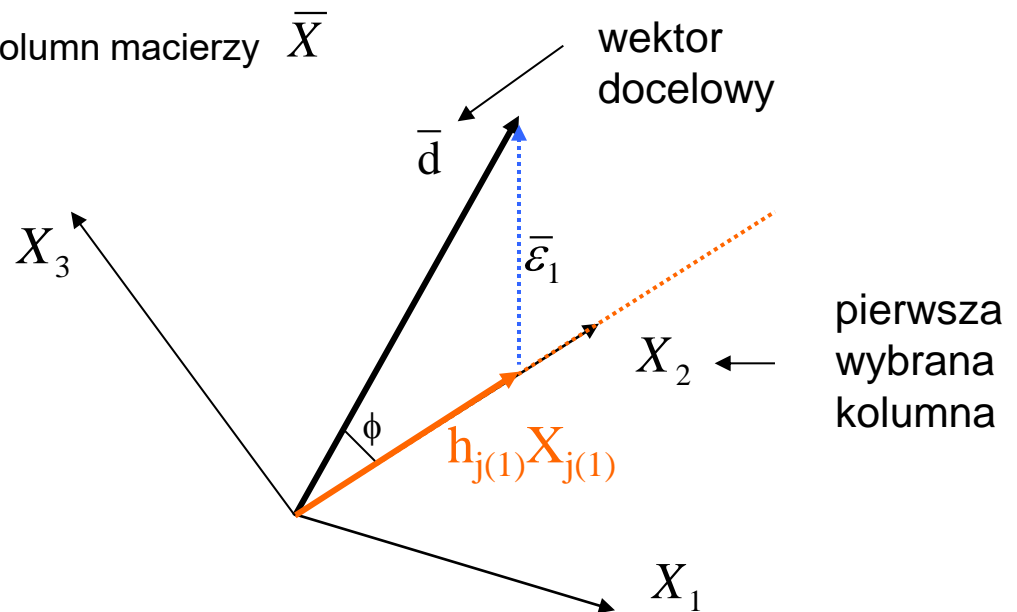
„Rzadki” model wektora docelowego  $\bar{d}$  : 
$$\bar{d}^* = \sum_{i=1}^K h_{j(i)} X_{j(i)}$$

jest kombinacją liniową  $K < M$  kolumn macierzy  $\bar{X}$

## Matching pursuit:

Kolumny wybieramy kolejno,  
stopniowo zmniejszając błąd

$$\bar{\varepsilon}_i = \bar{d} - \bar{d}_i^*$$



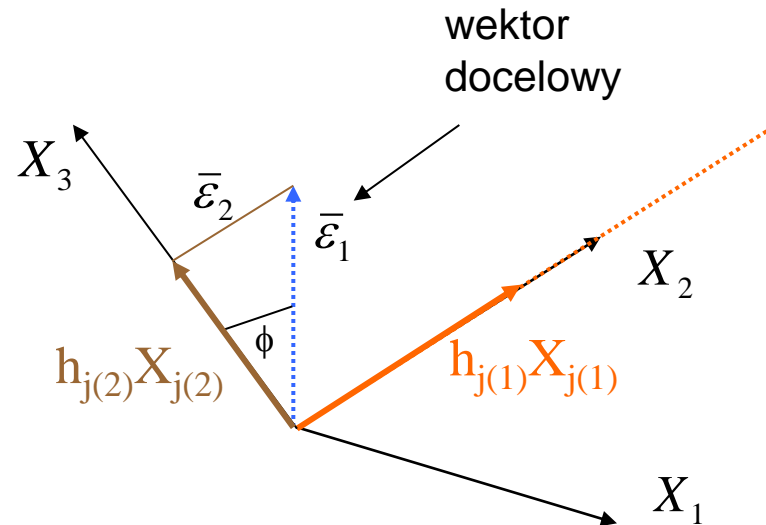
# „Rzadkie” modele – matching pursuit (MP)

Wybór drugiej kolumny:

## Matching pursuit:

Kolumny wybieramy kolejno,  
stopniowo zmniejszając błąd

$$\bar{\varepsilon} = \bar{d} - \bar{d}^*$$



# „Rzadkie” modele – OMP

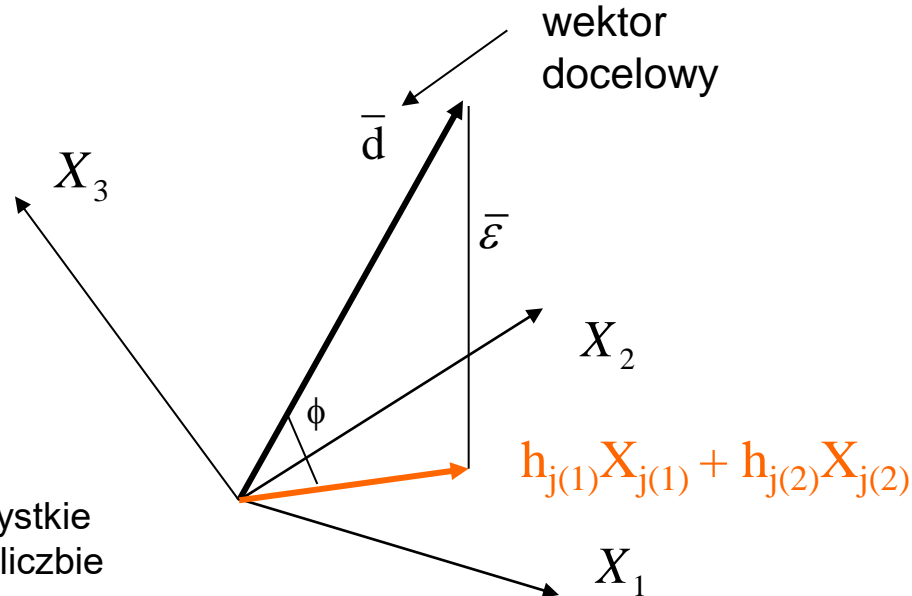
**Wada MP:** wolna zbieżność do wektora docelowego  
(odpowiednio mały błąd uzyskujemy przy dużym  $K$ , a więc model nie jest bardzo „rzadki”)

## Orthogonal MP (OMP):

rzutowanie wektora docelowego na podprzestrzeń rozpiętą na kolejno wybieranych kolumnach

OMP też nie daje gwarancji optymalnego rozwiązania, chyba że przeszukamy wszystkie możliwe podprzestrzenie w liczbie

$$\frac{M!}{K!(M-K)!}$$



# „Rzadkie” modele – min L1

W metodach LS minimalizujemy energię błędu wyrażoną normą euklidesową (norma L2):

$$\min_{\bar{h}} \|\bar{\varepsilon}\|_2^2 = \min_{\bar{h}} \|\bar{d} - \bar{X}\bar{h}\|_2^2 = \min_{\bar{h}} \sum_{i=1}^n \varepsilon_i^2$$

Dodajmy człon ograniczający normę L1 szukanej odpowiedzi impulsowej:

$$\|\bar{h}\|_1 = \sum_{j=0}^{M-1} |h_j|$$

Na przykład w ten sposób:

$$\min_{\bar{h}} \{ \lambda \|\bar{h}\|_1 + \|\bar{\varepsilon}\|_2^2 \} = \min_{\bar{h}} \{ \lambda \|\bar{h}\|_1 + \|\bar{d} - \bar{X}\bar{h}\|_2^2 \}$$

Takie zadanie ma analityczne rozwiązanie i zapewnia „rzadszą” odpowiedź impulsową niż metody LMS