Graz University of Technology

Adaptive Systems

Assignment 1

Autor: Ebner Thomas (0831246), Nöhmer Stefan (0830668)

Datum: Graz, 18. November 2011

Version.: alpha 1.0

1 Analytic Problem 1.1

1.0.1 (a)

Zuerst stellen wir eine Kostenfunktion $J(c)$, wie in den Problem-classes auf.

$$J(c) = E\{e[n] \cdot e[n]\} \tag{1.1}$$

Um das Minimum der Kostenfunktion zu ermitteln, bilden wir den Gradienten mit den Ableitungen nach den einzelnen Filterkoeffizienten $\mathbf{c}$.

$$\triangledown J(c) = E\{\triangledown e[n] \cdot e[n]\} \tag{1.2}$$

Setzt man die Beziehung $e[n] = d[n] - \mathbf{c}^T\mathbf{x}[n]$ in die Gleichung 1.2 ein, so erhält man:

$$\triangledown J(c) = E\{(\triangledown d[n] - \triangledown \mathbf{c}^T\mathbf{x}[n]) \cdot e[n]\} \tag{1.3}$$

$$\triangledown J(c) = E\{-\mathbf{x}[n] \cdot e[n]\} \tag{1.4}$$

Um nun das Minimum der Kostenfunktion zu erhalten müssen wir diesen Term 0 setzen:

$$\triangledown J(c) = E\{-\mathbf{x}[n] \cdot e[n]\} \stackrel{!}{=} 0 \tag{1.5}$$

Hiermit ist ersichtlich, dass jedes $x[n-k]$ orthogonal zu $e[n]$ ist.

1.0.2 (b)

Um diese Beziehung herzuleiten verwenden wir die Gleichung 1.5 und multiplizieren dise von links mit $\mathbf{c}^T$. Somit erhalten wir:

$$E\{-\mathbf{c}^T\mathbf{x}[n] \cdot e[n]\} \stackrel{!}{=} 0 \tag{1.6}$$

$$E\{-y[n] \cdot e[n]\} \stackrel{!}{=} 0 \tag{1.7}$$

Somit ist auch $y[n]$ orthogonal zu $e[n]$

1.0.3 (c)

Um den Wert für die Kostenfunktion bei optimalen Filterkoeffizienten zu berechnen, setzen wir einfach die Wiener Hopf Solution in die Gleichung für die Kostenfunktion (Gleichung 1.9).

$$J(\mathbf{c}) = E\{d^2[n] - 2d[n]\mathbf{c}^T\mathbf{x}[n] + \mathbf{c}^T\mathbf{x}[n]\mathbf{x}^T[n]\} \tag{1.8}$$

$$J(\mathbf{c}) = E\{d^2[n]\} - 2\mathbf{c}^T\mathbf{p} + \mathbf{c}^T\mathbf{R}_{xx} \tag{1.9}$$

Somit erhalten wir:

$$J(\mathbf{c}) = \sigma_d - 2\mathbf{p}^T\mathbf{R}_{xx}\mathbf{p} + \mathbf{p}^T\mathbf{R}_{xx}\mathbf{p} \tag{1.10}$$

$$J(\mathbf{c}) = \sigma_d - \mathbf{p}^T\mathbf{R}_{xx}\mathbf{p} \tag{1.11}$$

2 Analytic Problem 1.2

2.0.4 (a) Berechnung von $\underline{c}_{MSE}$

Zur Bestimmung von $\underline{c}_{MSE}$ muss zuerst eine Kostenfunktion $J(\underline{c})$ aufgestellt werden. Diese wird dann abgeleitet und eine Nullstelle gesucht (Suche nach dem Minimum). Dadurch erhält man $\underline{c}_{MSE}$, bei dem die Abweichung am geringsten ist. $x[n]$ und $\nu[n]$ sind dabei *jointly stationary.*

$$d[n] = \underline{h}^T \underline{x}[n] + \nu[n] \tag{2.1}$$

$$e[n] = d[n] - y[n] = \underline{h}^T \underline{x}[n] + \nu[n] - \underline{c}^T \underline{x}[n] \tag{2.2}$$

Die Kostenfunktion ergibt sich zu:

$$\begin{aligned}
J(\underline{c}) &= E\{e^2[n]\} = E\{(\underline{h}^T \underline{x}[n] + \nu[n] - \underline{c}^T \underline{x}[n])^2\} && (2.3)\\
&= E\{\underline{h}^T \underline{x}[n]\underline{x}^T[n]\underline{h} + \nu[n]\underline{h}^T \underline{x}[n] - \underbrace{\underline{h}^T \underline{x}[n]\underline{x}^T[n]\underline{c}}_{\underline{c}^T\underline{x}[n]\underline{x}^T[n]\underline{h}} + \nu[n]\underline{h}^T \underline{x}[n] + \nu^2[n] && (2.4)\\
&\quad -\nu[n]\underline{c}^T \underline{x}[n] - \underline{c}^T \underline{x}[n]\underline{x}^T[n]\underline{h} - \nu[n]\underline{c}^T \underline{x}[n] + \underline{c}^T \underline{x}[n]\underline{x}^T[n]\underline{c}\} && (2.5)\\
&= \underline{h}^T \underbrace{E\{\underline{x}[n]\underline{x}^T[n]\}}_{\underline{R}_{xx}} \underline{h} + 2\underline{h}^T \underbrace{E\{\nu[n]\underline{x}[n]\}}_{0} - 2\underline{c}^T \underbrace{E\{\underline{x}[n]\underline{x}^T[n]\}}_{\underline{R}_{xx}} \underline{h} + \underbrace{E\{\nu^2[n]\}}_{\sigma_\nu^2} && (2.6)\\
&\quad -2\underline{c}^T \underbrace{E\{\nu[n]\underline{x}[n]\}}_{0} + \underline{c}^T \underbrace{E\{\underline{x}[n]\underline{x}^T[n]\}}_{\underline{R}_{xx}} \underline{c} && (2.7)\\
&= \underline{h}^T \underline{R}_{xx}\underline{h} - 2\underline{c}^T \underline{R}_{xx}\underline{h} + \underline{c}^T \underline{R}_{xx}\underline{c} + \sigma_\mu^2 && (2.8)
\end{aligned}$$

Diese wird abgeleitet:

$$\bigtriangledown_{\underline{c}_{MSE}} J(\underline{c}_{MSE}) = 0 - 2\underline{R}_{xx}\underline{h} + 2\underline{R}_{xx}\underline{c}_{MSE} + 0 \overset{!}{=} 0 \tag{2.9}$$

Daraus ergibt sich:

$$\underline{R}_{xx}\underline{h} = \underline{R}_{xx}\underline{c}_{MSE} \Rightarrow \underline{c}_{MSE} = \underline{h} \tag{2.10}$$

Vorraussetzung ist dabei, dass $\underline{R}_{xx}^{-1}$ existiert.

2.0.5 (b) MMSE

$$J(\underline{c}_{MSE}) = \underline{h}^T \underline{R}_{xx}\underline{h} - 2\underline{h}^T \underline{R}_{xx}\underline{h} + \underline{h}^T \underline{R}_{xx}\underline{h} + \sigma_\mu^2 = \sigma_\mu^2 \tag{2.11}$$

2.0.6 (c) Bestimmung von $\underline{c}_{MSE}$ für bekanntes $\underline{h}$

Bei der Bestimmung von $\underline{c}_{MSE}$ müssen die 3 unterschiedlichen Fälle betrachtet werden: $N < N_h$, $N = N_h$ und $N > N_h$.

N=2

In diesem Fall muss die Berechnung erneut gemacht werden, diesmal aber mit aufgeteilten Vektoren:

$$\underline{x}'[n] = \begin{bmatrix} \underline{x}_a[n] \\ \underline{x}_b[n] \end{bmatrix}$$

Die Grösse des Vektors $\underline{x}[n]$ ist $(N_h, 1)$, Teilvektor $\underline{x}_a[n]$ ist $(N, 1)$, Teilvektor $\underline{x}_b[n]$ ist $(N_h - N, 1)$.

$$d[n] = \underbrace{\underline{h}^T}_{(1,N_h)} \underbrace{\underline{x}[n]}_{(N_h,1)} + \nu[n] \tag{2.12}$$

$$y[n] = \underbrace{\underline{c}^T}_{(1,N)} \underbrace{\underline{x}_a[n]}_{(N,1)} \tag{2.13}$$

$$e[n] = d[n] - y[n] = \underline{h}^T \underline{x}[n] + \nu[n] - \underline{c}^T \underline{x}_a[n] \tag{2.14}$$

Man kann jetzt neue Korrelationsmatrizen definieren:

$$E\{\underline{x}[n]\underline{x}^T[n]\} = E\{\begin{bmatrix} \underline{x}_a[n] \\ \underline{x}_b[n] \end{bmatrix} \begin{bmatrix} \underline{x}_a[n] & \underline{x}_b[n] \end{bmatrix}\} = E\{\begin{bmatrix} \underline{x}_a[n]\underline{x}_a^T[n] & \underline{x}_a[n]\underline{x}_b^T[n] \\ \underline{x}_b[n]\underline{x}_a^T[n] & \underline{x}_b[n]\underline{x}_b^T[n] \end{bmatrix}\} = \begin{bmatrix} \underline{R}_{aa} & \underline{R}_{ab} \\ \underline{R}_{ba} & \underline{R}_{bb} \end{bmatrix} \tag{2.15}$$

$$E\{\underline{x}[n]\underline{x}_a^T[n]\} = E\{\begin{bmatrix} \underline{x}_a[n] \\ \underline{x}_b[n] \end{bmatrix} \underline{x}_a[n]\} = E\{\begin{bmatrix} \underline{x}_a[n]\underline{x}_a^T[n] \\ \underline{x}_a[n]\underline{x}_a^T[n] \end{bmatrix}\} = \begin{bmatrix} \underline{R}_{aa} \\ \underline{R}_{ba} \end{bmatrix} \tag{2.16}$$

$$E\{\underline{x}_a[n]\underline{x}^T[n]\} = E\{\underline{x}_a[n] \begin{bmatrix} \underline{x}_a^T[n] & \underline{x}_b^T[n] \end{bmatrix}\} = E\{\begin{bmatrix} \underline{x}_a[n]\underline{x}_a^T[n] \\ \underline{x}_a[n]\underline{x}_b^T[n] \end{bmatrix}\} = \begin{bmatrix} \underline{R}_{aa} & \underline{R}_{ab} \end{bmatrix} \tag{2.17}$$

Die Kostenfunktion ergibt sich für dieses Beispiel zu:

$$\begin{aligned}
J(\underline{c}) &= E\{e^2[n]\} = E\{(\underline{h}^T \underline{x}[n] + \nu[n] - \underline{c}^T \underline{x}_a[n])^2\} && (2.18)\\
&= E\{\underline{h}^T \underline{x}[n]\underline{x}^T[n]\underline{h} + \nu[n]\underline{h}^T \underline{x}[n] - \underbrace{\underline{h}^T \underline{x}[n]\underline{x}_a^T[n]\underline{c}}_{\underline{c}^T \underline{x}_a[n]\underline{x}^T[n]\underline{h}} + \nu[n]\underline{h}^T \underline{x}[n] + \nu^2[n] && (2.19)\\
&\quad -\nu[n]\underline{c}^T \underline{x}_a[n] - \underline{c}^T \underline{x}_a[n]\underline{x}^T[n]\underline{h} - \nu[n]\underline{c}^T \underline{x}_a[n] + \underline{c}^T \underline{x}_a[n]\underline{x}_a^T[n]\underline{c}\} && (2.20)\\
&= \underline{h}^T E\{\underline{x}[n]\underline{x}^T[n]\}\underline{h} + 2\underline{h}^T \underbrace{E\{\nu[n]\underline{x}[n]\}}_{0} - 2\underline{c}^T E\{\underline{x}_a[n]\underline{x}^T[n]\}\underline{h} + \underbrace{E\{\nu^2[n]\}}_{\sigma_\nu^2} && (2.21)\\
&\quad -2\underline{c}^T \underbrace{E\{\nu[n]\underline{x}_a[n]\}}_{0} + \underline{c}^T E\{\underline{x}_a[n]\underline{x}_a^T[n]\}\underline{c} && (2.22)\\
&= \underline{h}^T \underline{R}_{xx}\underline{h} - 2\underline{c}^T [\underline{R}_{aa}\underline{R}_{ab}]\underline{h} + \underline{c}^T \underline{R}_{aa}\underline{c} + \sigma_\mu^2 && (2.23)
\end{aligned}$$

Die Kostenfunktion wird wieder abgeleitet und 0 gesetzt:

$$\underline{\nabla}_{\underline{c}_{MSE}} J(\underline{c}_{MSE}) = 0 + 0 - 2\begin{bmatrix} \underline{R}_{aa} & \underline{R}_{ab} \end{bmatrix} \underline{h} + 2\underline{R}_{xx}\underline{c}_{MSE} + 0 \stackrel{!}{=} 0 \tag{2.24}$$

Daraus ergibt sich:

$$\begin{bmatrix}\underline{R}_{aa} & \underline{R}_{ab}\end{bmatrix}\begin{bmatrix}\underline{h}_a \\ \underline{h}_b\end{bmatrix} = \underline{R}_{aa}\underline{c}_{MSE} \Rightarrow \underline{c}_{MSE} = \underline{R}_{aa}^{-1}(\underline{R}_{aa}\underline{h}_a + \underline{R}_{ab}\underline{h}_b) = \underline{h}_a\underline{R}_{aa}^{-1}\underline{R}_{ab}\underline{h}_b) \tag{2.25}$$

Da $x[n]$ ein *white, zero-mean* Signal ist, gilt $E\{x[n]x[n-k]\} = 0\forall k$, d.h. $\underline{R}_{xx}$ ist eine Diagonalmatrix, wobei alle Elemente in der Diagonale den Wert σ_x^2, also $\underline{R}_{xx} = \underline{I}\sigma_x^2$. Dadurch ergibt sich auch, dass $\underline{R}_{ab} = \underline{R}_{ba} = \underline{0}$. Dadurch ergibt sich:

$$\underline{c}_{MSE} = \underline{h}_a\underline{R}_{aa}^{-1}\underline{0}\underline{h}_b = \underline{h}_a = \begin{bmatrix}1 \\ -1\end{bmatrix} \tag{2.26}$$

Der MMSE berechnet sich zu:

$$\begin{aligned}
J(\underline{c}_{MSE}) &= \underline{h}^T\underline{R}_{xx}\underline{h} - 2\underline{h}_a^T\begin{bmatrix}\underline{R}_{aa} & \underline{R}_{ab}\end{bmatrix}\underline{h} + \underline{h}_a^T\underline{R}_{aa}\underline{h}_a + \sigma_\mu^2 && (2.27)\\
&= \begin{bmatrix}\underline{h}_a^T & \underline{h}_b^T\end{bmatrix}\begin{bmatrix}\underline{R}_{aa} & \underline{R}_{ab} \\ \underline{R}_{ba} & \underline{R}_{bb}\end{bmatrix}\begin{bmatrix}\underline{h}_a \\ \underline{h}_b\end{bmatrix} - 2\underline{h}_a^T\begin{bmatrix}\underline{R}_{aa} & \underline{R}_{ab}\end{bmatrix}\begin{bmatrix}\underline{h}_a \\ \underline{h}_b\end{bmatrix} + \underline{h}_a^T\underline{R}_{aa}\underline{h}_a + \sigma_\mu^2 && (2.28)\\
&= \underline{h}_a^T\underline{R}_{aa}\underline{h}_a + \underline{h}_a^T\underline{R}_{ab}\underline{h}_b + \underline{h}_b^T\underline{R}_{ba}\underline{h}_a + \underline{h}_b^T\underline{R}_{bb}\underline{h}_b - 2\underline{h}_a^T\underline{R}_{aa}\underline{h}_a && (2.29)\\
&\quad -2\underline{h}_a^T\underline{R}_{ab}\underline{h}_b + \underline{h}_a^T\underline{R}_{aa}\underline{h}_a + \sigma_\nu^2 && (2.30)\\
&= -\underline{h}_a^T\underline{R}_{ab}\underline{h}_b + \underline{h}_b^T\underline{R}_{ba}\underline{h}_a + \underline{h}_b^T\underline{R}_{bb}\underline{h}_b + \sigma_\nu^2 && (2.31)
\end{aligned}$$

Da $\underline{R}_{ab} = \underline{R}_{ba} = \underline{0}$ ergibt sich:

$$MMSE = -\underline{h}_a^T\underline{0}\underline{h}_b + \underline{h}_b^T\underline{0}\underline{h}_a + \underline{h}_b^T\underline{R}_{bb}\underline{h}_b + \sigma_\nu^2 = \underline{h}_b^T\underline{R}_{bb}\underline{h}_b + \sigma_\nu^2 \tag{2.32}$$

Mit $\underline{h}_b = 1$ (das letzte Element von $\underline{h}$ und $\underline{R}_{bb} = 1$ (das rechte untere Element von $\underline{R}_{xx}$, also $\sigma_x^2 = 1$) erhält man für den MMSE:

$$MMSE = 1 \cdot 1 \cdot 1 + 0.5 = 1.5 \tag{2.33}$$

Das adaptive System hat eine geringere Ordnung als das unbekannte System, es ist also “unterbestimmt”. Wie in den Problem Classes erklärt kann ein adaptives System einer geringeren Ordnung als das unbekannte System dieses nicht richtig abbilden, dadurch steigt der Fehler (und auch der MMSE).

N=3

Dies ist der einfache Fall mit $N = N_h$, die mit der Gleichung aus Aufgabe (a) berechnet werden kann.

$$\underline{c}_{MSE} = \underline{h} = \begin{bmatrix}1 \\ -1 \\ 1\end{bmatrix} \tag{2.34}$$

$$MMSE = \sigma_\nu^2 = 0.5 \tag{2.35}$$

In diesem Fall ist die Ordnung des adaptiven Systems gleich der Ordnung des unbekannten Systems. Es kann also gut abgebildet werden, und der Fehler wird minimiert. Der MMSE entspricht der Rauschleistung des Eingangsrauschens.

N=4

Dieser Fall (überbestimmt) ist auch einfach zu lösen. Wie in den Problem Classes gezeigt können die Vektoren mit 0 aufgefüllt werden:

$$\underline{h}' = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \underline{x}'[n] = \begin{bmatrix} \underline{x}[n] \\ 0 \end{bmatrix}$$

$$\underline{c}_{MSE} = \underline{h} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \tag{2.36}$$

$$MMSE = \sigma_\nu^2 = 0.5 \tag{2.37}$$

In diesem Fall ist die Ordnung des adaptiven Systems grösser als die Ordnung des unbekannten Systems. Deswegen kann es auch wieder gut abgebildet werden und der Fehler wird minimal (wie im Fall $N = N_h$).

2.0.7 (d) Berechnung der Korrelationsvektoren von ν

Aus der Definition von $r_{\nu x}$ erhält man:

$$r_{\nu x}[k] = E\{\nu[n]x[n-k]\} = E\{\frac{1}{2}(x[n]+x[n-1])x[n-k]\} = \frac{1}{2}E\{x[n]x[n-k]\}+\frac{1}{2}E\{x[n-1]x[n-k]\} \tag{2.38}$$

Da $x[n]$ ein *zero-mean, white gaussian noise* ist, gilt: $r_{xx}[0] = \sigma_x^2 = 1, r_{xx}[k] = 0 \forall k \neq 0$. Betrachten wir die möglichen Fälle:

$$k = 0 : r_{\nu x}[0] = \frac{1}{2}E\{x[n]x[n]\} + \frac{1}{2}E\{x[n-1]x[n]\} = \frac{1}{2}r_{xx}[0] + \frac{1}{2}r_{xx}[1] = \frac{1}{2} \tag{2.39}$$

$$k = 1 : r_{\nu x}[1] = \frac{1}{2}E\{x[n]x[n-1]\} + \frac{1}{2}E\{x[n-1]x[n-1]\} = \frac{1}{2}r_{xx}[1] + \frac{1}{2}r_{xx}[0] = \frac{1}{2} \tag{2.40}$$

$$k \geq 2 : r_{\nu x}[k] = \frac{1}{2}E\{x[n]x[n-k]\} + \frac{1}{2}E\{x[n-1]x[n-k]\} = \frac{1}{2}r_{xx}[k] + \frac{1}{2}r_{xx}[k-1] = 0 \tag{2.41}$$

Für $r_{\nu\nu}$ ergibt sich:

$$\begin{aligned} r_{\nu\nu}[k] &= E\{\nu[n]\nu[n-k]\} = E\{\frac{1}{2}(x[n] + x[n-1])\frac{1}{2}(x[n-k] + x[n-k-1])\} && (2.42) \\ &= \frac{1}{4}E\{x[n] + x[n-k]\} + \frac{1}{4}E\{x[n] + x[n-k-1]\} && (2.43) \\ &\quad +\frac{1}{4}E\{x[n-1] + x[n-k]\} + \frac{1}{4}E\{x[n-1] + x[n-k-1]\} && (2.44) \\ &= \frac{1}{4}r_{xx}[k] + \frac{1}{4}r_{xx}[k+1] + \frac{1}{4}r_{xx}[k-1] + \frac{1}{4}r_{xx}[k] && (2.45) \end{aligned}$$

Wir unterscheiden wieder die möglichen Fälle:

$$k = 0 : r_{\nu\nu}[0] = \frac{1}{4}r_{xx}[0] + \frac{1}{4}r_{xx}[1] + \frac{1}{4}r_{xx}[-1] + \frac{1}{4}r_{xx}[0] = \frac{1}{2} \tag{2.46}$$

$$k = 1 : r_{\nu\nu}[1] = \frac{1}{4}r_{xx}[1] + \frac{1}{4}r_{xx}[2] + \frac{1}{4}r_{xx}[0] + \frac{1}{4}r_{xx}[1] = \frac{1}{4} \tag{2.47}$$

$$k \geq 2 : r_{\nu\nu}[2] = \frac{1}{4}r_{xx}[k] + \frac{1}{4}r_{xx}[k+1] + \frac{1}{4}r_{xx}[k-1] + \frac{1}{4}r_{xx}[k] = 0 \tag{2.48}$$

Für die Varianz σ_ν^2 erhält man mit $\sigma_x^2 = 1$:

$$\begin{aligned}
\sigma_\nu^2 &= E\{\nu^2[n]\} = E\{\frac{1}{2}(x[n] + x[n-1])\frac{1}{2}(x[n] + x[n-1])\} && (2.49)\\
&= \frac{1}{4}E\{x^2[n] + 2x[n]x[n-1] + x^2[n-1]\} && (2.50)\\
&= \frac{1}{4}E\{x^2[n]\} + 2\frac{1}{4}E\{x[n]x[n-1]\} + \frac{1}{4}E\{x^2[n-1]\} && (2.51)\\
&= \frac{1}{4}\sigma_x^2 + \frac{1}{2}r_{xx}[1] + \frac{1}{4}\sigma_x^2 = \frac{1}{4}1 + \frac{1}{2}0 + \frac{1}{4}1 = \frac{1}{2} && (2.52)
\end{aligned}$$

2.0.8 (e) $\underline{c}_{MSE}$ für nicht unkorrelierte x und ν

Es müssen wieder neue Funktionen definiert werden:

$$d[n] = \underline{h}^T \underline{x}[n] + \frac{1}{2}(x[n] + x[n-1]) \tag{2.53}$$

$$y[n] = \underline{c}^T \underline{x}[n] \tag{2.54}$$

$$e[n] = d[n] - y[n] = \underline{h}^T \underline{x}[n] + \frac{1}{2}(x[n] + x[n-1]) - \underline{c}^T \underline{x}[n] \tag{2.55}$$

Für zukünftigen Gebrauch bestimmen wir $E\{(x[n] + x[n-1])\underline{x}[n]\}$ $(N = 3)$:

$$\begin{aligned}
E\{(x[n] + x[n-1])\underline{x}[n]\} &= E\{x[n]\underline{x}[n]\} + E\{x[n-1]\underline{x}[n]\} && (2.56)\\
&= E\{x[n]\begin{bmatrix} x[n] \\ x[n-1] \\ x[n-2] \end{bmatrix}\} + E\{x[n-1]\begin{bmatrix} x[n] \\ x[n-1] \\ x[n-2] \end{bmatrix}\} && (2.57)\\
&= \begin{bmatrix} \sigma_x^2 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \sigma_x^2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} =: \underline{r} && (2.58)
\end{aligned}$$

Damit wird wieder eine neue Kostenfunktion aufgestellt:

$$\begin{aligned}
J(\underline{c}) &= E\{e^2[n]\} = E\{(\underline{h}^T\underline{x}[n] + \frac{1}{2}(x[n]+x[n-1]) - \underline{c}^T\underline{x}[n])^2\} && (2.59)\\
&= E\{\underline{h}^T\underline{x}[n]\underline{x}^T[n]\underline{h} + \frac{1}{2}(x[n]+x[n-1])\underline{h}^T\underline{x}[n] - \underbrace{\underline{h}^T\underline{x}[n]\underline{x}^T[n]\underline{c}}_{\underline{c}^T\underline{x}[n]\underline{x}^T[n]\underline{h}} && (2.60)\\
&\quad +\frac{1}{2}(x[n]+x[n-1])\underline{h}^T\underline{x}[n] + \frac{1}{4}(x[n]+x[n-1])(x[n]+x[n-1]) && (2.61)\\
&\quad -\frac{1}{2}(x[n]+x[n-1])\underline{c}^T\underline{x}[n] - \underline{c}^T\underline{x}[n]\underline{x}^T[n]\underline{h} - \frac{1}{2}(x[n]+x[n-1])\underline{c}^T\underline{x}[n] && (2.62)\\
&\quad +\underline{c}^T\underline{x}[n]\underline{x}^T[n]\underline{c}\} && (2.63)\\
&= \underline{h}^T\underbrace{E\{\underline{x}[n]\underline{x}^T[n]\}}_{\underline{R}_{xx}}\underline{h} + \underline{h}^T\underbrace{E\{(x[n]+x[n-1])\underline{x}[n]\}}_{\underline{r}} - 2\underline{c}^T\underbrace{E\{\underline{x}[n]\underline{x}^T[n]\}}_{\underline{R}_{xx}}\underline{h} && (2.64)\\
&\quad +\underbrace{E\{\nu^2[n]\}}_{\sigma_\nu^2} - \underline{c}^T\underbrace{E\{(x[n]+x[n-1])\underline{x}[n]\}}_{0} + \underline{c}^T\underbrace{E\{\underline{x}[n]\underline{x}^T[n]\}}_{\underline{R}_{xx}}\underline{c} && (2.65)\\
&= \underline{h}^T\underline{R}_{xx}\underline{h} + \underline{h}^T\underline{r} - 2\underline{c}^T\underline{R}_{xx}\underline{h} - \underline{c}^T\underline{r} + \underline{c}^T\underline{R}_{xx}\underline{c} + \sigma_\mu^2 && (2.66)
\end{aligned}$$

Diese Kostenfunktion wird wieder abgeleitet und 0 gesetzt:

$$\bigtriangledown_{\underline{c}_{MSE}} J(\underline{c}_{MSE}) = 0 + 0 - 2\underline{R}_{xx}\underline{h} + 0 - \underline{r} + 2\underline{R}_{xx}\underline{c}_{MSE} \stackrel{!}{=} 0 \qquad (2.67)$$

Daraus ergibt sich (mit $\underline{r}' = \frac{1}{2}\underline{r}$):

$$2\underline{R}_{xx}\underline{h}\underline{r} = 2\underline{R}_{xx}\underline{c}_{MSE} \Rightarrow \underline{c}_{MSE} = \underline{R}_{xx}^{-1}(\underline{R}_{xx}\underline{h} * \frac{1}{2}\underline{r}) = \underline{h} + \underline{R}_{xx}^{-1}\underline{r}' \qquad (2.68)$$

Dadurch, dass ν und x nicht mehr unkorrelliert sind, verändert sich das Ergebnis! Der weitere Term, der hinzukommt, basiert auf σ_x^2, der Rauschleistung von x.
Nun bestimmen wir den MMSE. Dabei ist folgendes wichtig: da $x[n]$ *zero-mean, white gaussian noise* ist, ist $\underline{R}_{xx}$ eine Diagonalmatrix mit σ_x^2 in den Elementen der Diagonale. In diesem Fall ist also $\underline{R}_{xx}$ die Einheitsmatrix $\underline{I}$! Die Inverse $\underline{R}_{xx}^{-1}$ ist also auch die Einheitsmatrix!

$$\begin{aligned}
J(\underline{c}_{MSE}) &= \underline{h}^T\underline{R}_{xx}\underline{h} + \underline{h}^T\underline{r} - 2(\underline{h} + \underline{R}_{xx}^{-1}\underline{r}')^T\underline{R}_{xx}\underline{h} - (\underline{h} + \underline{R}_{xx}^{-1}\underline{r}')^T\underline{r} && (2.69)\\
&\quad +(\underline{h} + \underline{R}_{xx}^{-1}\underline{r}')^T\underline{R}_{xx}(\underline{h} + \underline{R}_{xx}^{-1}\underline{r}') + \sigma_\mu^2 && (2.70)\\
&= \underline{h}^T\underline{R}_{xx}\underline{h} + \underline{h}^T\underline{r} + (\underline{h} + \underline{R}_{xx}^{-1}\underline{r}')^T(-2\underline{R}_{xx}\underline{h} - \underline{r} + \underline{R}_{xx}(\underline{h} + \underline{R}_{xx}^{-1}\underline{r}')) + \sigma_\mu^2 && (2.71)\\
&= \underline{h}^T\underline{R}_{xx}\underline{h} + \underline{h}^T\underline{r} + (\underline{h} + \underline{R}_{xx}^{-1}\underline{r}')^T(-2\underline{R}_{xx}\underline{h} - \underline{r} + \underline{R}_{xx}\underline{h} + \underline{R}_{xx}\underline{R}_{xx}^{-1}\underline{r}')) + \sigma_\mu^2 && (2.72)\\
&= \underline{h}^T\underline{R}_{xx}\underline{h} + 2\underline{h}^T\underline{r}' + (\underline{h} + \underline{R}_{xx}^{-1}\underline{r}')^T(-2\underline{R}_{xx}\underline{h} - 2\underline{r}' + \underline{R}_{xx}\underline{h} + \underline{r}')) + \sigma_\mu^2 && (2.73)\\
&= \underline{h}^T\underline{R}_{xx}\underline{h} + 2\underline{h}^T\underline{r}' - (\underline{h} + \underline{R}_{xx}^{-1}\underline{r}')^T(\underline{R}_{xx}\underline{h} + \underline{r}') + \sigma_\mu^2 && (2.74)\\
&= \underline{h}^T\underline{R}_{xx}\underline{h} + 2\underline{h}^T\underline{r}' - (\underline{h}^T\underline{R}_{xx}\underline{h} + \underline{h}^T\underline{r}' + \underline{r}'^T\underline{R}_{xx}^{-1^T}\underline{R}_{xx}\underline{h} + \underline{r}'^T\underline{R}_{xx}^{-1^T}\underline{r}') + \sigma_\mu^2 && (2.75)\\
&= -\underline{h}^T\underline{r}' - \underbrace{\underline{r}'^T\underline{h}}_{\underline{h}^T\underline{r}'} + 2\underline{h}^T\underline{r}' - \underline{r}'^T\underline{R}_{xx}^{-1^T}\underline{r}' + \sigma_\mu^2 && (2.76)\\
&= -\underline{r}'^T\underline{R}_{xx}^{-1^T}\underline{r}' + \sigma_\mu^2 && (2.77)
\end{aligned}$$

Da $\underline{R}_{xx}$ symmetrisch ist (Töplitz-Struktur), erhält man mit $(\underline{R}_{xx}^{-1})^T = (\underline{R}_{xx}^T)^{-1} = \underline{R}_{xx}^{-1}$:

$$MMSE = J(\underline{c}_{MSE}) = \sigma_\mu^2 - \underline{r}'^T\underline{R}_{xx}^{-1^T}\underline{r}' \qquad (2.78)$$

Da $\underline{R}_{xx}$ wie bereits erwähnt eine Diagonalmatrix mit 1 in den Diagonalelementen ist, ist $\underline{R}_{xx}^{-1}$ ebenfalls eine Diagonalmatrix mit 1 in den Diagonalelementen (Einheitsmatrix). Mit $\underline{r}' = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}$ erhält man für den MMSE:

$$MMSE = 0.5 - \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix} = 0.5 - \frac{1}{2} = 0 \tag{2.79}$$

Man erkennt, dass sich der Fehler auf 0 verringert hat! Durch den bekannten Zusammenhang zwischen ν und x lässt sich das System exakt an die Bedingungen anpassen.
Berechnet man nun den MMSE mit $\underline{c} = \underline{h}$, so ergibt sich:

$$MMSE = J(\underline{c} = \underline{h}) = \underline{h}^T \underline{R}_{xx} \underline{h} + \underline{h}^T \underline{r} - 2\underline{h}^T \underline{R}_{xx} \underline{h} - \underline{h}^T \underline{r} + \underline{h}^T \underline{R}_{xx} \underline{h} + \sigma_\mu^2 = \sigma_\mu^2 = 0.5 \tag{2.80}$$

Bei der Verwendung dieser Anpassung entspricht der Fehler wieder der Rauschleistung σ_ν^2.

3 MATLAB Problem 1.3

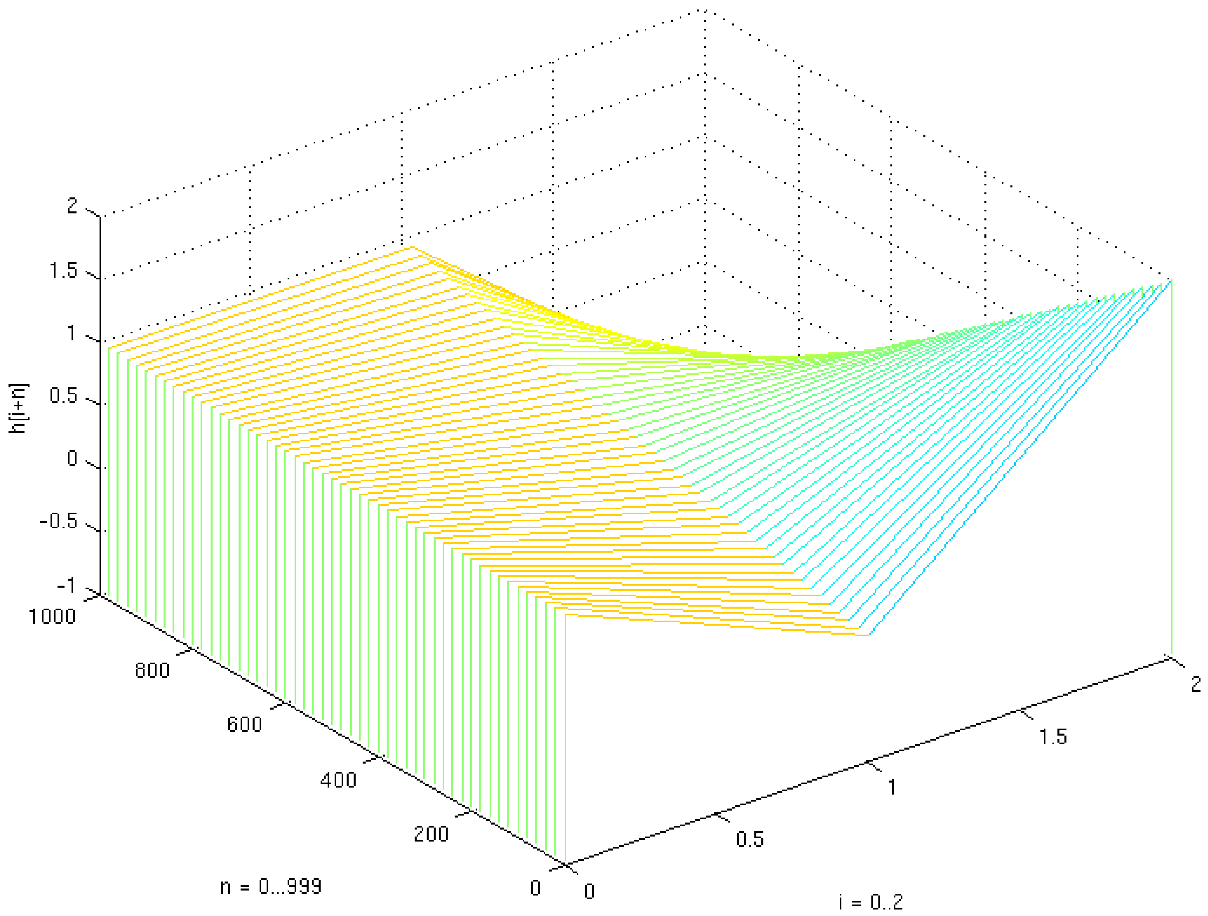

Abbildung 3.1: Wasserfall Darstellung der Impulsantwort

3.0.9 (c) Adaptiver Filter ohne Rauschen

In der Abbildung 3.2 bzw. 3.3 wurden die Koeffizienten des "unbekannten" Systems mit den Koeffizienten des Adaptiven Systems verglichen.
Dabei kann man erkennen, dass bei einem größerem M das unbekannte System besser angenähert werden kann. Die Koeffizienten haben wie in den Abbildungen ersichtlich eine kleinere Varianz. Das ganze ist natürlich logisch, da längere Blöcke verarbeitet werden.
Nachteil eines größeren M ist ein etwas größerer Berechnungsaufwand. Ein weiterer Nachteil welcher zwar hier nicht ersichtlich ist, dass die Filterkoeffizienten bei schnelleren Änderungen von $\mathbf{h}[n]$ nicht mitkommen.

3.0.10 (d) Adaptiver Filter mit Weißem Rauschen

In den Abbildungen 3.4 bzw. 3.5 ist zu erkennen, dass das Hinzufügen von Weißem Rauschen keinen wesentlichen Unterschied bringt. Die Abweichung von den idealen Koeffizienten ist natürlich etwas größer.

3.0.11 (e) Adaptiver Filter mit vom Eingang abhängigem $\nu[n]$

In den Abbildungen 3.6 bzw. 3.7 ist zu erkennen, dass die durch das Adaptive Filter angenäherten Koeffizienten einen Offset zu den idealen Koeffizienten haben. Dies liegt daran, dass $\nu[n]$ abhängig vom Eingang $x[n]$ ist.
Betrachtet man die Gleichung:

$$\nu[n] = 1/2 \cdot (x[n] + x[n-1]) \tag{3.1}$$

so fällt auf, dass $\nu[n]$ aus einem Filter mit den Koeffizienten $[0.5; 0.5; 0]$ aus $x[n]$ erzeugt werden kann. Dies spiegelt sich auch in den Koeffizienten des Adpativen Filters wieder. Somit lässt sich der "Offset" der Koeffizienten in den Abbildungen 3.6 bzw. 3.7 erklären.

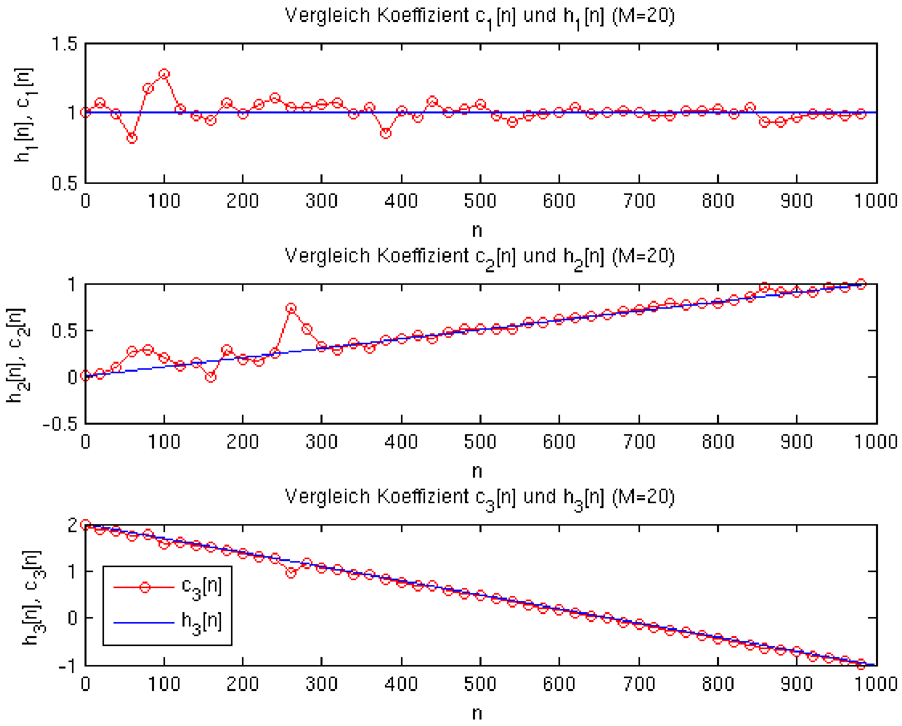

Abbildung 3.2: Koeffizienten des Filters mit M=20, $v[n] = 0$

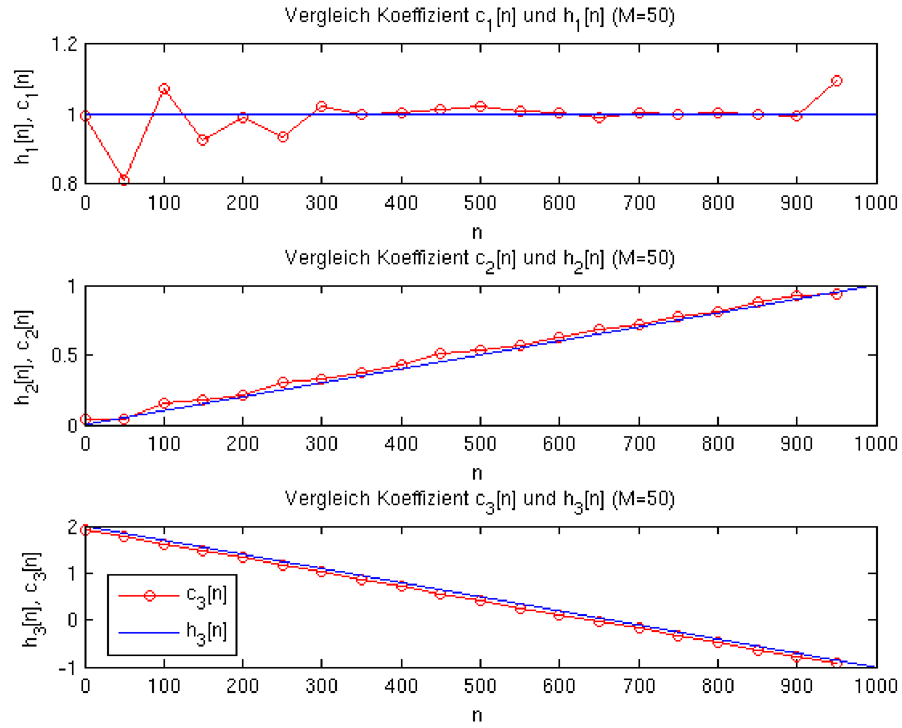

Abbildung 3.3: Koeffizienten des Filters mit M=50, $v[n] = 0$

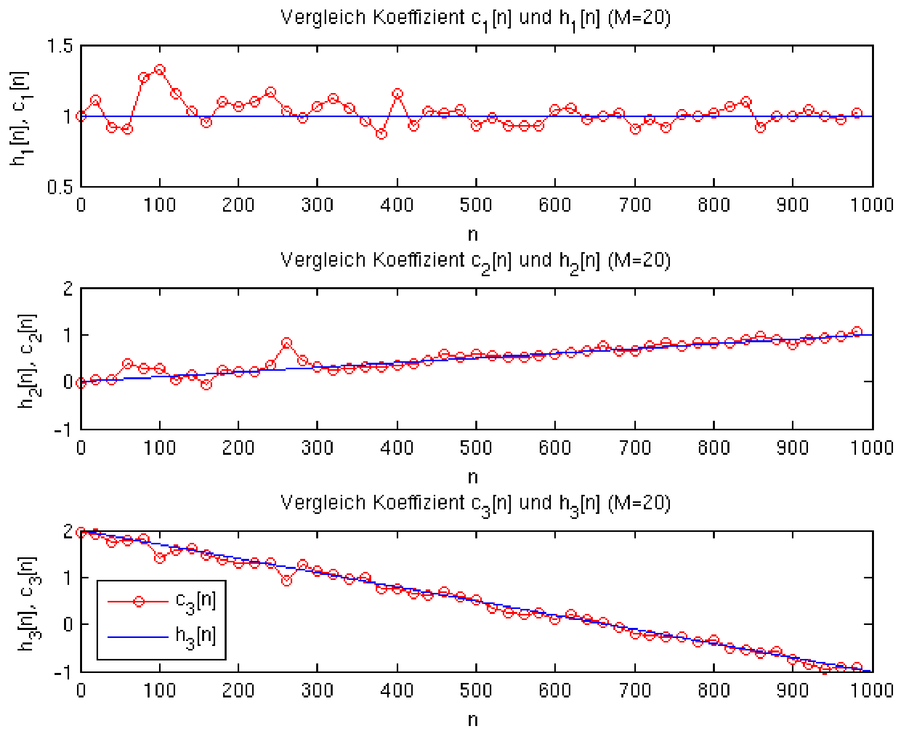

Abbildung 3.4: Koeffizienten des Filters mit M=20, v[n] = zero mean white noise with Variance = 0.5

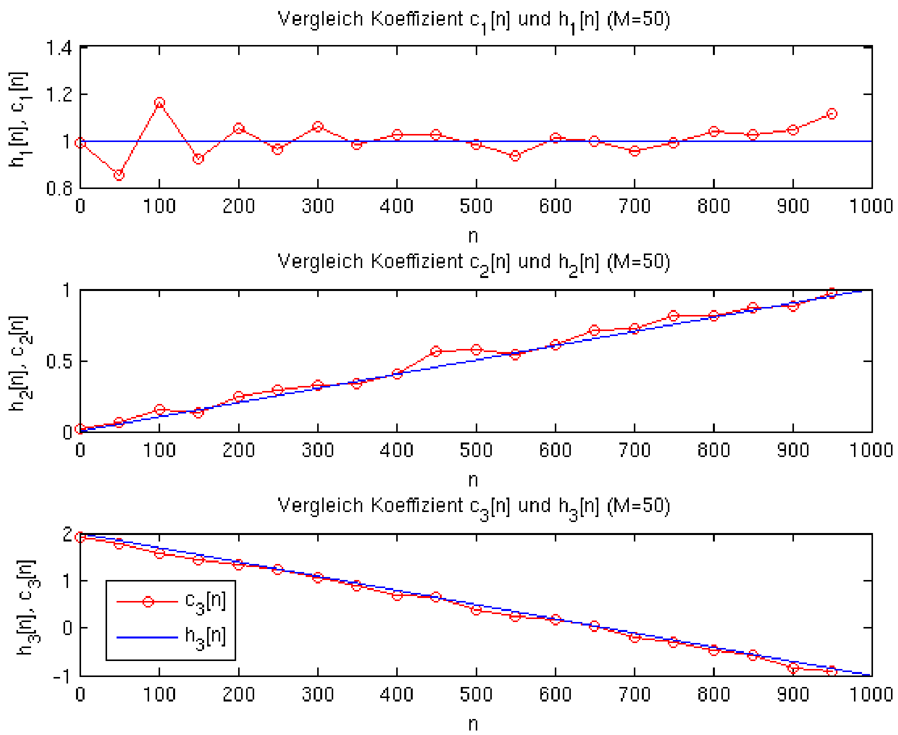

Abbildung 3.5: Koeffizienten des Filters mit M=50, v[n] = zero mean white noise with Variance = 0.5

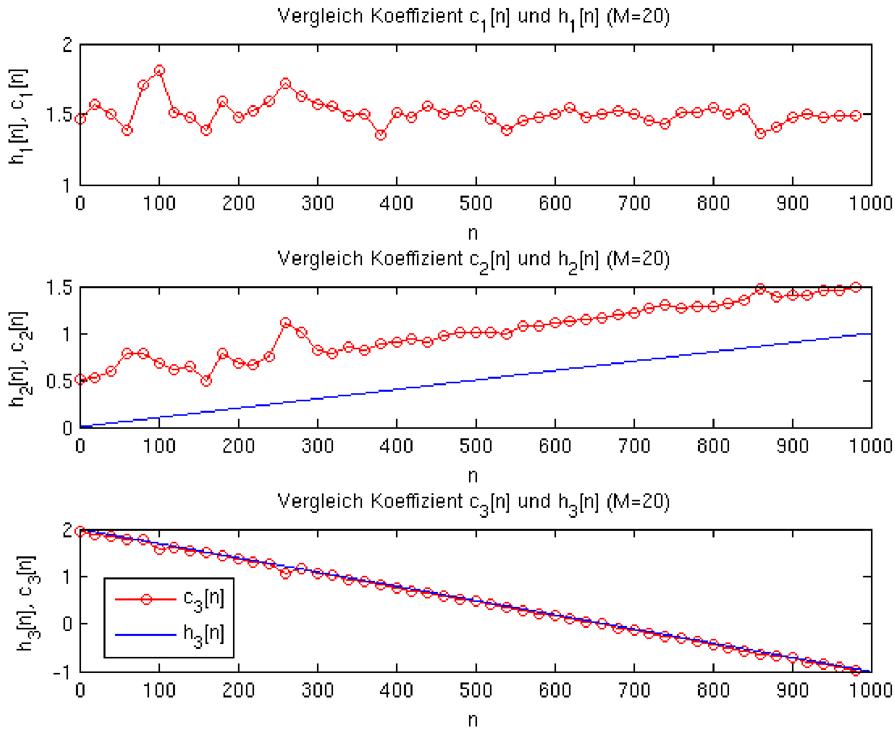

Abbildung 3.6: Koeffizienten des Filters mit M=20, $v[n] = 0.5 \cdot (x[n] + x[n-1]$

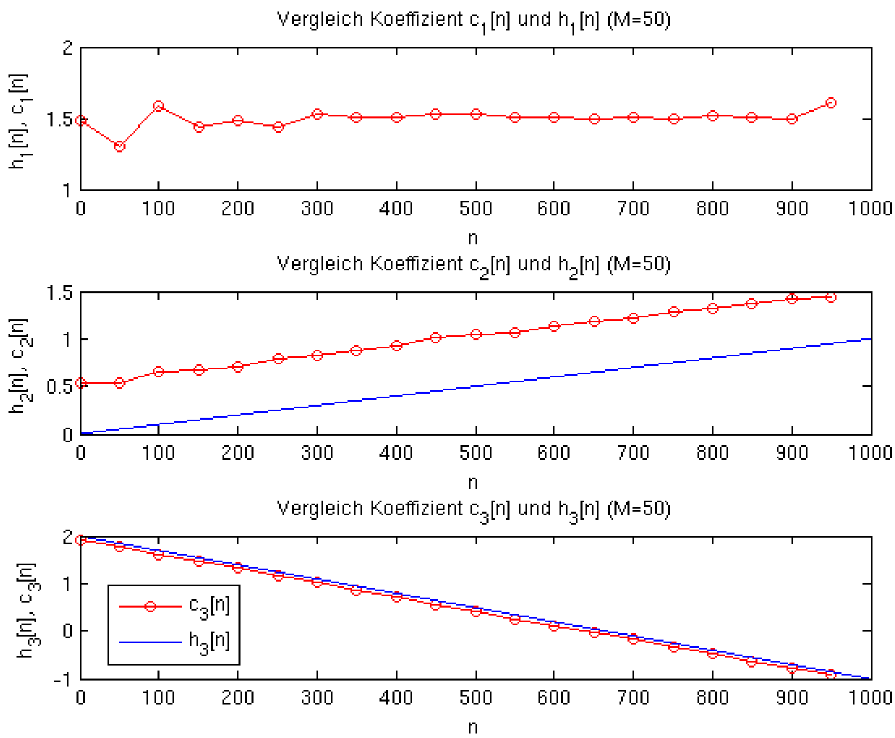

Abbildung 3.7: Koeffizienten des Filters mit M=50, $v[n] = 0.5 \cdot (x[n] + x[n-1]$

4 Listings

4.1 LS-Filter

```
function c = ls_filter( x, d, N)
% x ... input signal
% d ... desired output signal (of same length as x)
% N ... number of filter coefficients

x = x(:);
d = d(:);


M = length(x);
warning off;
X = toeplitz(x, zeros(N,1));
warning on;

%Moore-Penrose Inverse:
c = X \ d;

end
```

4.2 Unknown System

```
function [ y,h ] = unknownsystem(x)
%SYSTEM Summary of this function goes here
%   Detailed explanation goes here

x = x(:);

N = length(x);
M = 3;

warning off;
X = toeplitz(x, zeros(3,1));
warning on;

y = zeros(N + M - 1, 1);


h = zeros(3, N);

for n = 1:N
    h(:,n) = [1; 0.001*n; 2-0.003*n];

    y(n) = X(n,:) * h(:,n);
end


end
```

4.3 Gesamtskript 13

```
clear;
close all;

% 1.3 (b): Generieren des Wasserfall-Plots
```

```
N = 1000;

x = diag(ones(1,N));

y = zeros(N,3);

for n = 1:N
    [temp,~] = unknownsystem(x(n,:));
    y(n,:) = temp(n:n+2);
end


figure;

[mx, my] = meshgrid(0:2, 0:20:999);
waterfall(mx, my, y(1:20:1000, :));
xlabel('i = 0..2');
ylabel('n = 0...999');
zlabel('h[i+n]');



%%
% 1.3 (c): kein Rauschen

%generate input signal (white noise) with variance 1:
x = randn(1000,1);

[d_without_noise,h] = unknownsystem(x);

d = d_without_noise;
calc_c_and_plot;

%%
% 1.3 (d): Rauschen mit Varianz=0.5

v = 0.25*randn(1002,1);
d = d_without_noise + v;
calc_c_and_plot;

%%
% 1.3 (e): Rauschen: v[n] = 0.5(x[n] + x[n-1])

v(1) = 0;
for i = 2:1000
    v(i) = 0.5*(x(i)+x(i-1));
end
d = d_without_noise + v;
calc_c_and_plot;
```

4.4 Skript zum Berechnen und Plotten der Koeffizienten

```
N = 3;


%segment size 20:
M = 20;

c_20 = zeros(3,ceil(1000/M));

for i = 1:ceil(1000/M);
    range = (1+M*(i-1)):(M*(i));
    c_20(:,i) = ls_filter(x(range), d(range), N);
end


%segment size 50:
M = 50;

c_50 = zeros(3,ceil(1000/M));
```

```

for i = 1:ceil(1000/M);
    range = [1+M*(i-1):M*(i)];
    c_50(:,i) = ls_filter(x(range), d(range), N);
end

%Plotten der Koeffizienten fÃ¼r M=20
M = 20;
figure;
subplot(3,1,1);
plot(0:M:999, c_20(1,:), 'r-o');
hold on;
plot(0:999, h(1,:), 'b- ');
ylabel('h_1[n], c_1[n]');
xlabel('n');
title('Vergleich Koeffizient c_1[n] und h_1[n] (M=20)');

subplot(3,1,2);
plot(0:M:999, c_20(2,:), 'r-o');
hold on;
plot(0:999, h(2,:), 'b- ');
ylabel('h_2[n], c_2[n]');
xlabel('n');
title('Vergleich Koeffizient c_2[n] und h_2[n] (M=20)');

subplot(3,1,3);
plot(0:M:999, c_20(3,:), 'r-o');
hold on;
plot(0:999, h(3,:), 'b- ');
ylabel('h_3[n], c_3[n]');
xlabel('n');
title('Vergleich Koeffizient c_3[n] und h_3[n] (M=20)');
legend('c_3[n]', 'h_3[n]');


%Plotten der Koeffizienten fÃ¼r M=50
M = 50;
figure;
subplot(3,1,1);
plot(0:M:999, c_50(1,:), 'r-o');
hold on;
plot(0:999, h(1,:), 'b- ');
ylabel('h_1[n], c_1[n]');
xlabel('n');
title('Vergleich Koeffizient c_1[n] und h_1[n] (M=50)');

subplot(3,1,2);
plot(0:M:999, c_50(2,:), 'r-o');
hold on;
plot(0:999, h(2,:), 'b- ');
ylabel('h_2[n], c_2[n]');
xlabel('n');
title('Vergleich Koeffizient c_2[n] und h_2[n] (M=50)');

subplot(3,1,3);
plot(0:M:999, c_50(3,:), 'r-o');
hold on;
plot(0:999, h(3,:), 'b- ');
ylabel('h_3[n], c_3[n]');
xlabel('n');
title('Vergleich Koeffizient c_3[n] und h_3[n] (M=50)');
legend('c_3[n]', 'h_3[n]');
```