Digital Signal Processing

Jiří Málek

Part I

Optimal filter design - motivation (Least square error filter)

Optimal filters

- Until now we discussed the **frequency selective filters**
- Those are designed to have a specific shape of the frequency response (e.g., low-pass filter)
- The design does not require any training data, the filter is fully specified by the *filter specifications*

- The **optimal filters** follow a different design philosophy
- Their frequency response is designed by optimization of some criterion (e.g., least squares), evaluated on a set of *training signals*
- Properties of the optimal filters are thus given by the training signals; the design cannot proceed without them
- In general, these filters do not have frequency selective character
- Let us explain the optimal filter design using the example of *Target cancellation filters*

Part II

Target Cancellation Filters

Target Cancellation Filters I

Let us consider the following task:

- Two simultaneously active sound sources, recorded via a binaural microphone array
- Sound of each source should be recovered from the mixture separately
- ASSUMPTION 1: One source is fixed (two speaking people, one has to sit/stand, the other can be moving)
- ASSUMPTION 2: For some short time interval, the fixed source is active alone

- The fixed source can be attenuated using the **target cancellation filter** (an example of an optimal filter).
- The moving source can be removed using an **adaptive filter** (details are outside the scope of the Lecture).

Target Cancellation Filters II

Formal problem description:

- Stereo recording with simultaneously active fixed source $s[n]$ and moving source $y[n]$ given by

$$\begin{aligned} x_{\mathrm{L}}[n] &= \{h_{\mathrm{L}} * s\}[n] + y_{\mathrm{L}}[n], \\ x_{\mathrm{R}}[n] &= \{h_{\mathrm{R}} * s\}[n] + y_{\mathrm{R}}[n]. \end{aligned} \tag{1}$$

  - $x_{\mathrm{L}}[n]$, $x_{\mathrm{R}}[n]$ ... two channels of the recordings
  - $y_{\mathrm{L}}[n]$, $y_{\mathrm{R}}[n]$ ... signal $y[n]$ at left and right microphone
  - $h_{\mathrm{L}}[n]$, $h_{\mathrm{R}}[n]$ ... acoustic impulse responses binding $s[n]$ to its image on left and right microphone

- NOTE: $\{h_{\mathrm{L}} * s\}[n]$ is an alternative way to denote convolution $\{h_{\mathrm{L}}[n] * s[n]\}$, which is sometimes used in the literature
- It emphasizes that we consider the nth sample of the sequence given by convolution of $h_{\mathrm{L}}[n]$ and $s[n]$.

Target Cancellation Filter (CF):

- CF blocks signal arriving from one direction in the environment
- CF is *time invariant* (i.e. LTI system), the target cannot change its position (ASSUMPTION 1).
- CF is a two-input single-output filter (Multi Input Single Output, MISO).

MISO filters in general:

- A set of single channel filters $g_i[n], i = 1 \dots I$ (for CF, $I = 2$)
- *Application:* Convolution of impulse response $g_i[n]$ with ith channel of the input (for $i = 1 \dots I$) and summation of all the output signals sample-wise.

Adaptive filter: changes its inner parameters (impulse/frequency response) in time.

Target Cancellation Filters IV

- CF can be designed as two filters: a general $g[n]$ (to be designed via least squares) and simple delay $g_2[n] = -\delta[n]$.
- Coefficients of g are selected to fulfill

$$\{g * h_\mathrm{L}\}[n] = h_\mathrm{R}[n], \tag{2}$$

i.e., the response of the left sensor (with respect to source $s[n]$) is filtered to be equal with the response of the right sensor.

- The filter $g[n]$ is thus given through the room impulse responses $h_L[n], h_R[n]$, which depend on the acoustic properties of the environment (reverberation) and the location of the fixed source
- The application of the CF results into output

$$\begin{aligned} v[n] &= \{g * x_\mathrm{L}\}[n] + \{g_2 * x_\mathrm{R}\}[n] \\ &= \{g * x_\mathrm{L}\}[n] - x_\mathrm{R}[n] \\ &= \{g * h_\mathrm{L} * s\}[n] + \{g * y_\mathrm{L}\}[n] - \{h_\mathrm{R} * s\}[n] - y_\mathrm{R}[n] \\ &= \{h_\mathrm{R} * s\}[n] + \{g * y_\mathrm{L}\}[n] - \{h_\mathrm{R} * s\}[n] - y_\mathrm{R}[n] \\ &= \{g * y_\mathrm{L}\}[n] - y_\mathrm{R}[n], \end{aligned} \tag{3}$$

which does not contain $s[n]$, while y_L and y_R are passed through.

- Signal $v[n]$ thus represents our estimate of the moving source $y[n]$.

Target Cancellation Filters V

How to compute g in practice:

- Filter g is computed using some interval $(n = N_1, \ldots, N_2)$ within $x_{\mathrm{L}}[n]$ and $x_{\mathrm{R}}[n]$, where only the fixed source is active (Assumption 2).
- If $y_{\mathrm{L}}(n) = y_{\mathrm{R}}(n) = 0$, then g is given as a solution ot a set of equations

$$g = \arg\min_{g} \sum_{n=N_1}^{N_2} \left| \{g * x_{\mathrm{L}} - x_{\mathrm{R}}\}[n] \right|^2. \tag{4}$$

- This is a classical problem in signal processing, it is a **least squares design** of an optimal filter
- The filter is thus given through the training signals $x_L[n], x_R[n]$.
- More precisely: by definition, $g[n]$ is given solely by impulse responses $h_L[n], h_R[n]$. In practice though, the estimate is also influenced by the signal $s[n]$, because $x_L[n] = \{h_L * s\}[n]$ a $x_R[n] = \{h_R * s\}[n]$.
- The training interval can be short (about 1 s of a signal).

Target Cancellation Filters VI

Adaptive filtering:

- Using the estimate of $v[n]$, we can also estimate the target $s[n]$ using the adaptive filtering.
- The adaptive (Wiener) filter is given in the spectral domain as

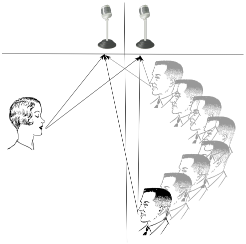
$$W[k,\ell] = \frac{|X[k,\ell]|^2}{|X[k,\ell]|^2 + \tau|V[k,\ell]|^2}. \tag{5}$$

  - $X[k,\ell]$, $V[k,\ell]$ ... short-time Discrete Fourier Transform (STFT) of signals $x_{\mathrm{L}}[n]$, $v[n]$
  - k ... index of a spectral bin
  - ℓ ... time index
  - τ ... free parameter (separation/distortion trade-off, classical Wiener $\tau = 1$)
- The STFT representation of target $s[n]$ is then given by

$$\widehat{S}[k,\ell] = W[k,\ell]X[k,\ell]. \tag{6}$$

- This is a variant of the *spectral masking*. We discussed another variant of this technique in the lecture about *spectral thresholding*.

Example:



- **Moving Source:** right half-space, variable distance to mics 0.5 m - 1.2 m, estimated using CF.
- **Fixed source:** left half-space, distance 1 m, estimated using adaptive filtering.

Original stereo recording:

Mixture channel 1

Mixture channel 2

Source estimates:

Moving source (via CF)

Fixed source (via masking)

Part III

Least square error optimal filters

Least square error optimal filter I

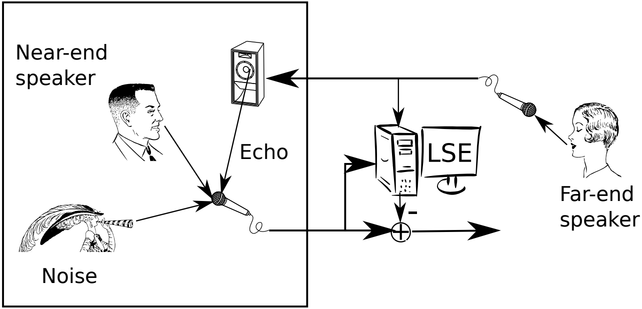
Least square error optimal filter: (LSE)

- The design minimizes least square difference between filter output and the desired target signal
- Filter design is completely dependent on data, frequency response is not known a priori
- *Beware of confusing naming:* Least square error optimal filter is a completely different system compared to least square frequency selective filter (Lecture 12, given by filter specifications)

- Many applications - prediction, dereverberation, system identification, ..
- .. target (directional signal) cancellation filters; minimizes difference between channels of recording originating from a binaural microphone array
- ..echo cancellation, attenuation of sound repetitions arising during duplex communication (hands-free, conference rooms)

Echo cancellation

- Duplex communication: sound emanating from a loudspeaker is captured by a microphone and send back
- This creates an unpleasant repetition of the original sound
- Some sounds (clicks) can even amplify with each pass through the loop
- In practice, the filter need to be successively adapted to changes in the environment (e.g., location of microphone/loudspeaker); adaptive LSE variants are used (e.g., Recursive Least Squares - RLS)



Least square error optimal filter II

- Let us continue with variables defined in the previous section. Then, the task of the LSE filter is to process the signal x_{L} (defined on interval $n \in [N_1, N_2]$) by filter g (of order N_g), such that the output signal

$$\hat{x}_{\mathrm{R}}[n] = \sum_{k=0}^{N_g} g[k] x_{\mathrm{L}}[n-k] \tag{7}$$

  was as similar as possible to a target signal x_{R} in the least square sense.

- Let us define the error signal

$$e[n] = x_{\mathrm{R}}[n] - \hat{x}_{\mathrm{R}}[n], \tag{8}$$

  which should have on a given signal interval as small energy as possible, i.e., let us minimize the criterion

$$J = \sum_{n=N_1}^{N_2} e^2[n]. \tag{9}$$

- Let us introduce a vector notation

$$\mathbf{x}_\mathrm{L}[n] = \begin{bmatrix} x_\mathrm{L}[n] \\ x_\mathrm{L}[n-1] \\ \vdots \\ x_\mathrm{L}[n-N_g+1] \end{bmatrix}, \mathbf{x}_\mathrm{R} = \begin{bmatrix} x_\mathrm{R}[N_1] \\ \vdots \\ x_\mathrm{R}[N_2] \end{bmatrix}, \hat{\mathbf{x}}_\mathrm{R} = \begin{bmatrix} \hat{x}_\mathrm{R}[N_1] \\ \vdots \\ \hat{x}_\mathrm{R}[N_2] \end{bmatrix}, \tag{10}$$

$$\mathbf{e} = \begin{bmatrix} e[N_1] \\ \vdots \\ e[N_2] \end{bmatrix}, \mathbf{g} = \begin{bmatrix} g[0] \\ \vdots \\ g[N_g] \end{bmatrix}. \tag{11}$$

- Convolution $\hat{x}_\mathrm{R} = g * x_\mathrm{L}$ can then be written as dot product

$$\hat{x}_\mathrm{R}[n] = \sum_{k=0}^{N_g} g[k] x_\mathrm{L}[n-k] = \mathbf{x}_\mathrm{L}^T[n]\mathbf{g} \tag{12}$$

Least square error optimal filter IV

- Error signal: $\mathbf{e} = \mathbf{x}_\mathrm{R} - \hat{\mathbf{x}}_\mathrm{R}$.
- Vector $\hat{\mathbf{x}}_\mathrm{R}$ can be also written as

$$\hat{\mathbf{x}}_\mathrm{R} = \begin{bmatrix} \mathbf{x}_\mathrm{L}^T[N_1]\mathbf{g} \\ \vdots \\ \mathbf{x}_\mathrm{L}^T[N_2]\mathbf{g} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_\mathrm{L}^T[N_1] \\ \vdots \\ \mathbf{x}_\mathrm{L}^T[N_2] \end{bmatrix} \cdot \mathbf{g} = \mathbf{A} \cdot \mathbf{g}. \tag{13}$$

- Martix $\mathbf{A}$ with dimensions $(N_2 - N_1 + 1) \times N_g$ is therefore

$$\mathbf{A} = \begin{bmatrix} x[N_1] & x[N_1 - 1] & \dots & x[N_1 - N_b + 1] \\ x[N_1 + 1] & x[N_1] & \dots & x[N_1 - N_b + 2] \\ \vdots & \vdots & \vdots & \vdots \\ x[N_2] & x[N_2 - 1] & \dots & x[N_2 - N_b + 1] \end{bmatrix}. \tag{14}$$

- Error signal: $\mathbf{e} = \mathbf{x}_\mathrm{R} - \hat{\mathbf{x}}_\mathrm{R} = \mathbf{x}_\mathrm{R} - \mathbf{A}\mathbf{g}$.

Least square error optimal filter V

- The task is to find such filter $\mathbf{g}$, such that the criterion measuring the energy of the error signal

$$J = \sum_{n=N_1}^{N_2} e^2[n]. \tag{15}$$

is as small as possible.

$$\begin{aligned} J &= \mathbf{e}^T\mathbf{e} = (\mathbf{x}_{\mathrm{R}} - \mathbf{A}\mathbf{g})^T(\mathbf{x}_{\mathrm{R}} - \mathbf{A}\mathbf{g}) \\ &= (\mathbf{x}_{\mathrm{R}}^T - \mathbf{g}^T\mathbf{A}^T)(\mathbf{x}_{\mathrm{R}} - \mathbf{A}\mathbf{g}) \\ &= \mathbf{x}_{\mathrm{R}}^T\mathbf{x}_{\mathrm{R}} - \mathbf{g}^T\mathbf{A}^T\mathbf{x}_{\mathrm{R}} - \mathbf{x}_{\mathrm{R}}^T\mathbf{A}\mathbf{g} + \mathbf{g}^T\mathbf{A}^T\mathbf{A}\mathbf{g} \\ &= \mathbf{x}_{\mathrm{R}}^T\mathbf{x}_{\mathrm{R}} - 2\mathbf{g}^T\mathbf{A}^T\mathbf{x}_{\mathrm{R}} + \mathbf{g}^T\mathbf{A}^T\mathbf{A}\mathbf{g}. \end{aligned} \tag{16}$$

- Global minimum is obtained by derivative of J by vector $\mathbf{g}$

$$\frac{dJ}{d\mathbf{g}} = \begin{bmatrix} \frac{dJ}{g[0]} \\ \vdots \\ \frac{dJ}{g[N_g]} \end{bmatrix}. \tag{17}$$

Least square error optimal filter VI

- By minimization of the criterion we obtain

$$\frac{dJ}{d\mathbf{g}} = -2\mathbf{A}^T\mathbf{x}_{\mathrm{R}} + 2\mathbf{A}^T\mathbf{A}\mathbf{g} \tag{18}$$

- After setting the derivative equal to zero vector we obtain

$$\mathbf{A}^T\mathbf{A}\mathbf{g} = \mathbf{A}^T\mathbf{x}_{\mathrm{R}} \tag{19}$$

- The analytic formula for the optimum filter is therefore

$$\mathbf{g} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{x}_{\mathrm{R}}. \tag{20}$$

- Matrix $\mathbf{A}^T\mathbf{A}$ of the set of linear equations (19) is toeplitz symmetric.
- Computation of $\mathbf{g}$ can be fastened considerably using Levinson-Durbinovy recursion (complexity $O(N_g^2)$)
- Classical approach to such solution is the Gauss elimination with complexity $O(N_g^3)$.

Thank you for attention!