Adaptive Filtering

Digital Signal Processing

Notes

Contents

Notes

Adaptive Filtering

Using filters that adapt or change depending on the signal characteristics.

Notes

Applications

- Electroencephalography
  - Eye movements can create noise bigger than the signal of interest and vary with time
- Digital spread spectrum communication
  - Noise or other signals may interfere at a particular band of frequencies that may vary over time
- High frequency digital data communication
  - over a communication channel with limited bandwidth

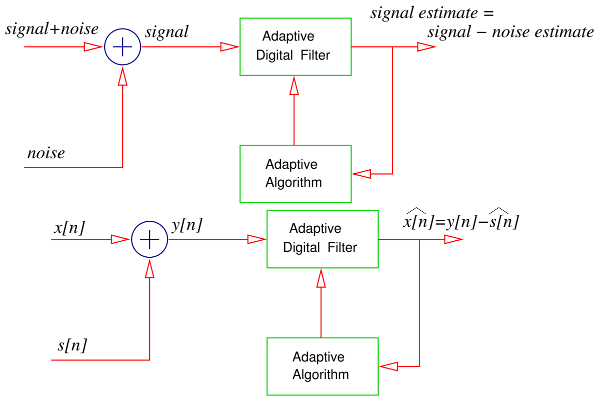
Notes

Adaptive Filters

- Variable filter characteristics, adapting to changing conditions
- Spectral overlap between signal and noise
- Unknown noise band or varies with time

Notes

Noise Cancellation



Notes

Some Definitions

- **Expectation** is the **mean** (a type of **average**). The discrete expectation from $n = 0$ to N for a signal $a[n]$ is:

$$\mathbb{E}\left[a[n]\right] = \frac{1}{N}\sum_{n=0}^{N} a[n].$$

- **Total power** is the **expectation** of the **square** of a signal:

$$\mathbb{E}\left[a^2[n]\right].$$

- **Signal to Noise** Ratio (SNR) is:

$$\frac{\mathbb{E}\left[a^2[n]\right]}{\mathbb{E}\left[b^2[n]\right]}$$

  where $b[n]$ is a **noise** signal.
- If two signals ($a[n]$ and $b[n]$) are **not correlated** then

$$\mathbb{E}\left[a[n]b[n]\right] = 0.$$

Notes

Adaptive Noise Canceller *Example*

[1] **Q.** The output of a noise canceller system is given by:

$$\hat{x}[n] = y[n] - \hat{s}[n] = x[n] + s[n] - \hat{s}[n]$$

Show that minimizing the total power at output maximises the output signal to noise ratio.
A. Previous slide showed that **power** and **signal to noise ratio** use **expectations and squares**.
So we need to find the **expectations** of the **squares** of the left and right sides of the equation.

[1] Widrow *et al.* 1975 and Ifeacher and Jervis 1993 pp. 546

Notes

Adaptive Noise Canceller *Example cont'd.*

1. Square both sides:

$$\begin{aligned}\hat{x}^2[n] &= (x[n] + s[n] - \hat{s}[n])^2\\ &= (x[n] + (s[n] - \hat{s}[n]))^2\\ &= x^2[n] + 2x[n](s[n] - \hat{s}[n]) + (s[n] - \hat{s}[n])^2\end{aligned}$$

2. Take expectations of both sides:

$$\begin{aligned}\mathbb{E}\left[\hat{x}^2[n]\right] &= \mathbb{E}\left[x^2[n] + 2x[n](s[n] - \hat{s}[n]) + (s[n] - \hat{s}[n])^2\right]\\ &= \mathbb{E}\left[x^2[n]\right] + \mathbb{E}\left[2x[n](s[n] - \hat{s}[n])\right] + \mathbb{E}\left[(s[n] - \hat{s}[n])^2\right]\end{aligned}$$

3. The input signal $x[n]$ and noise source $s[n]$ are **not correlated**,

$$\mathbb{E}\left[2x[n](s[n] - \hat{s}[n])\right] = 0$$

   (see definitions slide) so

$$\begin{aligned}\mathbb{E}\left[\hat{x}^2[n]\right] &= \mathbb{E}\left[x^2[n]\right] + 0 + \mathbb{E}\left[(s[n] - \hat{s}[n])^2\right]\\ &= \mathbb{E}\left[x^2[n]\right] + \mathbb{E}\left[(s[n] - \hat{s}[n])^2\right]\end{aligned}$$

Notes

Adaptive Noise Canceller *Example cont'd.*

$$\mathbb{E}\left[\hat{x}^2[n]\right] = \mathbb{E}\left[x^2[n]\right] + \mathbb{E}\left[(s[n]-\hat{s}[n])^2\right]$$

- Total power at output:
$$\mathbb{E}\left[\hat{x}^2[n]\right]$$
- Signal power at output:
$$\mathbb{E}\left[x^2[n]\right]$$
- (Remaining) noise power at output:
$$\mathbb{E}\left[(s[n]-\hat{s}[n])^2\right]$$
- Output signal to noise rato:
$$\frac{\mathbb{E}\left[x^2[n]\right]}{\mathbb{E}\left[(s[n]-\hat{s}[n])^2\right]}$$

Signal to noise ratio increases if $\mathbb{E}\left[(s[n]-\hat{s}[n])^2\right] \to 0$ which **minimizes**:

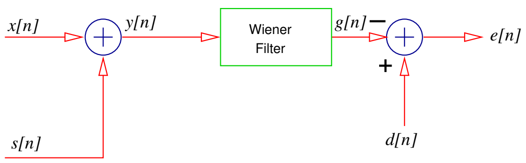
$$\min\left(\mathbb{E}\left[\hat{x}^2[n]\right]\right) = \mathbb{E}\left[x^2[n]\right] + \min\left(\mathbb{E}\left[(s[n]-\hat{s}[n])^2\right]\right).$$

So minimizing the total power at output increases the signal to noise ratio.

Notes

Wiener Filter

Many **adaptive algorithms** := Discrete **Wiener filter** approximations



- Input signal=true signal + noise **or** $y[n] = x[n] + s[n]$
- **Wiener filtered signal:** $g[n]$
- **Error:** $e[n] = d[n] - g[n]$
- **Signal of interest:** $d[n]$

Objective:
minimise **Mean Square** of the Error (**MSE**)

- **Error:** $e[n] = d[n] - g[n]$
- **MSE:** $\mathbb{E}\left[(e[n])^2\right] = \mathbb{E}\left[(d[n]-g[n])^2\right]$

Notes

Wiener Filter

- Wiener FIR filtered signal $g[n] = \sum_{m=0}^{M} w[m]y[n-m] = \mathbf{w}^T Y$

  Convolution of signal $y[n]$ *with Wiener filter coefficients* $w[n]$

  where $\mathbf{y} = \begin{pmatrix} y[0] \\ y[1] \\ \dots \\ y[m] \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} w[n] \\ w[n-1] \\ \dots \\ w[n-m] \end{pmatrix}$.
- So error:
$$e[n] = d[n] - g[n] = d[n] - \mathbf{w}^{\mathrm{T}}\mathbf{y}.$$
- Square both sides of error:
$$\begin{aligned} e[n]^2 &= \left(d[n] - \mathbf{w}^{\mathrm{T}}\mathbf{y}\right)^2 = d^2[n] - 2d[n]\mathbf{w}^{\mathrm{T}}\mathbf{y} + \left(\mathbf{w}^{\mathrm{T}}\mathbf{y}\right)^2 \\ &= d^2[n] - 2d[n]\mathbf{w}^{\mathrm{T}}\mathbf{y} + \left(\mathbf{w}^{\mathrm{T}}\mathbf{y}\right)\left(\mathbf{w}^{\mathrm{T}}\mathbf{y}\right) \end{aligned}$$

Notes

Wiener Filter

- As $\mathbf{w}^{\mathsf{T}}\mathbf{y} = \mathbf{y}^{\mathsf{T}}\mathbf{w}$ then

$$\begin{aligned} e[n]^2 &= d^2[n] - 2\mathbf{w}^{\mathsf{T}}\mathbf{y}d[n] + \left(\mathbf{w}^{\mathsf{T}}\mathbf{y}\right)\left(\mathbf{y}^{\mathsf{T}}\mathbf{w}\right) \\ &= d^2[n] - 2\mathbf{w}^{\mathsf{T}}\mathbf{y}d[n] + \mathbf{w}^{\mathsf{T}}\mathbf{y}\mathbf{y}^{\mathsf{T}}\mathbf{w} \end{aligned}$$

- Taking expectations (average):

$$\begin{aligned} \mathbb{E}\left[e[n]^2\right] &= \mathbb{E}\left[d^2[n] - 2\mathbf{w}^{\mathsf{T}}\mathbf{y}d[n] + \mathbf{w}^{\mathsf{T}}\mathbf{y}\mathbf{y}^{\mathsf{T}}\mathbf{w}\right] \\ &= \mathbb{E}\left[d^2[n]\right] - \mathbb{E}\left[2\mathbf{w}^{\mathsf{T}}\mathbf{y}d[n]\right] + \mathbb{E}\left[\mathbf{w}^{\mathsf{T}}\mathbf{y}\mathbf{y}^{\mathsf{T}}\mathbf{w}\right] \\ &= \mathbb{E}\left[d^2[n]\right] - 2\mathbf{w}^{\mathsf{T}}\mathbb{E}\left[\mathbf{y}d[n]\right] + \mathbf{w}^{\mathsf{T}}\mathbb{E}\left[\mathbf{y}\mathbf{y}^{T}\right]\mathbf{w} \end{aligned}$$

- Let $\mathbf{R}_y = \mathbb{E}\left[\mathbf{y}\mathbf{y}^T\right]$ and $\mathbf{r}_{dy} = \mathbb{E}\left[\mathbf{y}d[n]\right]$ then

$$\mathbb{E}\left[e[n]^2\right] = \mathbb{E}\left[d^2[n]\right] - 2\mathbf{w}^{\mathsf{T}}\mathbf{r}_{dy} + \mathbf{w}^{\mathsf{T}}\mathbf{R}_y\mathbf{w}$$

where $\mathbf{R}_y$ is the autocorrelation matrix and $\mathbf{r}_{dy}$ is the cross-correlation.

Notes

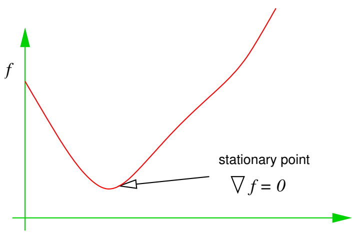
Wiener Filter

Mean Square Error:

$$\mathbb{E}\left[e[n]^2\right] = \mathbb{E}\left[d^2[n]\right] - 2\mathbf{w}^{\mathsf{T}}\mathbf{r}_{dy} + \mathbf{w}^{\mathsf{T}}\mathbf{R}_y\mathbf{w}$$

Q. How to minimize error?

A. A function has a stationary point where the gradient is zero:



Therefore take the gradient and set to zero:
to find **optimal** FIR filter weights.

Notes

Wiener Filter

Derivative of Mean Square Error:

$$\begin{aligned} \frac{\mathrm{d}\left(\mathbb{E}\left[e[n]^2\right]\right)}{\mathrm{d}\mathbf{w}} &= \frac{\mathrm{d}\left(\mathbb{E}\left[d^2[n]\right] - 2\mathbf{w}^{\mathsf{T}}\mathbf{r}_{dy} + \mathbf{w}^{\mathsf{T}}\mathbf{R}_y\mathbf{w}\right)}{\mathrm{d}\mathbf{w}} \\ &= \frac{\mathrm{d}\left(\mathbb{E}\left[d^2[n]\right]\right)}{\mathrm{d}\mathbf{w}} - \frac{\mathrm{d}\left(2\mathbf{w}^{\mathsf{T}}\mathbf{r}_{dy}\right)}{\mathrm{d}\mathbf{w}} + \frac{\mathrm{d}\left(\mathbf{w}^{\mathsf{T}}\mathbf{R}_y\mathbf{w}\right)}{\mathrm{d}\mathbf{w}} \\ &= 0 - 2\mathbf{r}_{dy} + 2\mathbf{R}_y\mathbf{w} \end{aligned}$$

Making derivative equal to zero to find the stationary point (minimum):

$$\begin{aligned} -2\mathbf{r}_{dy} + 2\mathbf{R}_y\mathbf{w} &= 0 \\ 2\mathbf{R}_y\mathbf{w} &= 2\mathbf{r}_{dy} \\ \mathbf{w} &= \mathbf{R}_y^{-1}\mathbf{r}_{dy} \end{aligned}$$

So optimal weights given by inverse of autocorrelation matrix multiplied with the cross correlation.

Notes

Wiener Filter Optimisation Algorithms

Result known as Wiener-Hopf equation:

$$\mathbf{R}_y \mathbf{w} = \mathbf{r}_{dy} \qquad (1)$$

Solution can be found by:

$$\mathbf{w} = \mathbf{R}_y^{-1} \mathbf{r}_{dy}$$

But **Levinson-Durbin algorithm** can be used to solve eqn (1) directly.
Alternatively there are the:

- **Least Mean Squares** (LMS) adaptive algorithm or
- **Recursive Least Squares** (RLS) algorithm

Notes

Least Mean Squares Adaptive Algorithm

[2]

- Wiener filter coefficients calculated iteratively
- Uses steepest descent algorithm:

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \mu \frac{\mathrm{d}\mathbb{E}\left[e[n]\right]^2}{\mathrm{d}\mathbf{w}}$$

- However $\frac{\mathrm{d}\mathbb{E}[e[n]]^2}{\mathrm{d}\mathbf{w}}$ is computationally expensive (as before)
- Wiener LMS algorithm uses:

$$\mathbf{w}_{n+1} = \mathbf{w}_n - 2\mu e[n] \mathbf{y}$$

  which is much faster to calculate.
- Does not depend on calculating $\mathbf{r}_{dy}$, $\mathbf{R}_y^{-1}$ *or* $\mathbf{R}_y$.

[2] Widrow *et al.* 1975

Notes

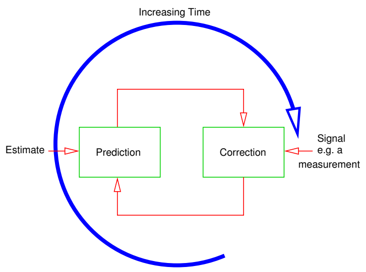
Kalman Filter

Wiener filter originally designed for signals that do not change with time (*stationary* signals).

- Can be adapted using *e.g.* LMS adaptive algorithm.

Kalman filters were invented for signals that change with time (*non-stationary* signals).

Filtering a **non-stationary signal**:



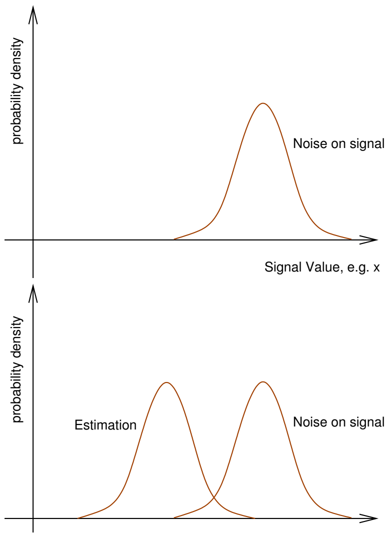
Notes

Kalman Filter

Kalman filter combines:

- Measurement
- and Prediction

in an algorithm to reduce the overall error when estimating a time varying signal (so probably not a periodic signal).



Notes

Lecture Summary

- Adaptive filtering uses filters that can adapt to a signal.
- Mean Square Error (MSE) is used in optimization algorithms to find suitable filter coefficients.
- Wiener filter is one technique.
- Wiener filter was designed for stationary signals.
- Kalman filtering is an example of a technique to estimate non-stationary signals.

Notes

Notes