Self-Orthogonalizing Adaptive Filters (con't)

Interpretation as a Rotation of the Error Surface

The MSE surface is given by

$$\begin{aligned} J(n) &= \sigma_\delta^2 - \boldsymbol{\pi}^H \mathbf{P}^{-1}\boldsymbol{\pi} + [\hat{\boldsymbol{\omega}}(\nu) - \mathbf{P}^{-1}\boldsymbol{\pi}]^H \mathbf{P}[\hat{\boldsymbol{\omega}}(\nu) - \mathbf{P}^{-1}\boldsymbol{\pi}] \\ &= \vartheta_{\mu\iota\nu} + \varepsilon^H(\nu)\mathbf{P}\,\varepsilon(\nu) \end{aligned}$$

For the transformed case we have

$$\begin{aligned} J(n) &= \sigma_\delta^2 - \boldsymbol{\pi}^H \mathbf{P}^{-1}\boldsymbol{\pi} + [\hat{\boldsymbol{\omega}}(\nu) - \mathbf{P}^{-1}\boldsymbol{\pi}]^T \boldsymbol{\Theta}\boldsymbol{\Theta}^T \mathbf{P}\boldsymbol{\Theta}\boldsymbol{\Theta}^T[\hat{\boldsymbol{\omega}}(\nu) - \mathbf{P}^{-1}\boldsymbol{\pi}] \\ &= \vartheta_{\mu\iota\nu} + \left[\hat{\boldsymbol{\omega}}(\nu)' - \boldsymbol{\omega}_o'\right]^T \Lambda \left[\hat{\boldsymbol{\omega}}(\nu)' - \boldsymbol{\omega}_o'\right] \\ &= \vartheta_{\mu\iota\nu} + \varepsilon^H(\nu)' \Lambda \varepsilon(\nu)' \end{aligned}$$

where

$$\varepsilon(n)' = \hat{\mathbf{w}}(n)' - \mathbf{w}_o'$$

and $\hat{\mathbf{w}}(n)'$, $\mathbf{w}_o'$ represent the coefficients of the transform-domain adaptive filter, transformed Wiener filter, respectively. The effect of applying KLT to the input signal is to rotate the error surface.

Figures 4.4 and 4.5 (Diniz)

We note that the eccentricity of the MSE surface remains unchanged and therefore the eigenvalue spread is unchanged by the transformation. Therefore no improvement in convergence rate is expected.

However, if in addition each element of the transformed output (filter coefficients) is normalized by its respective eigenvalue (pseudo power normalization), the distance between points where equal-error contours meet coefficient axes and the origin are equalized. As a result, a reduction in eigenvalue spread is expected.

Figure 4.6 (Diniz)

Therefore combining the KLT transform on the input signal vector and subsequent normalization on the correction yields an error surface with circular contours.

Figures 4.7 and 4.8 (Diniz)

$$\hat{\mathbf{w}}(n+1)' = \hat{\boldsymbol{\omega}}(\nu)' + \alpha\Lambda^{-1}\boldsymbol{\varpi}(\nu)\varepsilon(\nu)$$

Practical Considerations

The KLT is a signal-dependent transform. The implementation of which requires estimation of the correlation matrix, diagonalization of this matrix, and the construction of the required basis vectors ($\mathbf{q}_1$, ..., $\mathbf{q}_M$) for transformation. This effort make KLT impractical for real-time applications.

Alternatively, we choose a unitary transform that is close to the KLT of the particular input signal. By close we mean that both transforms perform nearly the same rotation of the MSE surface. Our transform would thus be pre-determined and *not* signal-dependent.

A Markov process is a stochastic process whose past has no influence on the future if its present is specified. This means the following:

if $t_{n-1} < t_n$ then

$$P\{\mathbf{x}(t_n) \le \xi_\nu | \boldsymbol{\xi}(\tau), \ \tau \le \tau_{\nu-1}\} = P\{\boldsymbol{\xi}(\tau_\nu) \le \xi_\nu | \boldsymbol{\xi}(\tau_{\nu-1})\}$$

For 1st-order Markov processes (correlations given by $r(n) = \alpha^{\nu}$) as well as speech signals the Discrete-Cosine Transform (DCT) is a good approximation for the KLT, i.e. the coefficients of the DCT provide an approximation to **Q**. Additionally, fast algorithms for the DCT exist.

We thus will use the DCT to compute an approximation to **v**(n) and we will estimate the eigenvalues of **R** to compute Λ^{-1}. The approximation and estimations will give rise to our DCT-LMS algorithm.

Discrete-Cosine Transform

The *M*-pt DCT pair is given by

$$U(m) = \kappa_\mu \sum_{\nu=0}^{M-1} u(\nu) \cos\left[\frac{(2\nu+1)\mu\pi}{2M}\right], \ 0 \le \mu \le M-1$$

$$u(\nu) = \frac{2}{M} \sum_{\mu=0}^{M-1} \kappa_\mu Y(\mu) \cos\left[\frac{(2\nu+1)\mu\pi}{2M}\right], \ 0 \le \nu \le M-1$$

where

$$k_m = \begin{cases} 1/\sqrt{2}, & \mu = 0 \\ 1, & 1 \le \mu \le M-1 \end{cases}.$$

Figure 10.3.