Kalman Filter as the Unifying Basis for RLS Filters

The bulk of this lecture comes from

Ali H. Sayed and Thomas Kailath, "A state-space approach to adaptive RLS filtering," *IEEE Signal Processing Magazine*, vol. 11, no. 3, pp. 18-60, July 1994.

Please note that there have been updates and corrections to the above paper posted on various FTP sites.

The Kalman filter provides a unifying framework for the derivation of those linear adaptive filters which constitute the family of RLS filters.

The key question is: given a Kalman filter based on a stochastic model, how do we derive the corresponding version of the RLS filter based on a deterministic model?

We list the two algorithms again

Recursive Least Squares	*Discrete Kalman Filter*
$\mathbf{k}(n) = \dfrac{\lambda^{-1}\mathbf{\Pi}(\nu-1)\mathbf{\upsilon}(\nu)}{1+\lambda^{-1}\mathbf{\upsilon}^H(\nu)\mathbf{\Pi}(\nu-1)\mathbf{\upsilon}(\nu)}$	$\hat{\mathbf{x}}(n\|n-1) = \mathbf{A}(\nu-1)\hat{\boldsymbol{\xi}}(\nu-1\|\nu-1)$
$\xi(n) = d(n) - \hat{\mathbf{w}}^H(n-1)\mathbf{u}(n)$	$\mathbf{P}(n\|n-1) = \mathbf{A}(n-1)\mathbf{P}(n-1\|n-1)\mathbf{A}^H(n-1) + \mathbf{Q}_w(n)$
$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}(n-1) + \mathbf{k}(n)\xi^*(n)$	$\mathbf{K}(n) = \mathbf{\Pi}(\nu\|\nu-1)\mathbf{X}^H(\nu)[\mathbf{X}(\nu)\mathbf{\Pi}(\nu\|\nu-1)\mathbf{X}^H(\nu) + \mathbf{\Theta}_{\varpi}(\nu)]^{-1}$
$\mathbf{P}(n) = \lambda^{-1}[\mathbf{I} - \mathbf{k}(n)\mathbf{u}^H(n)]\mathbf{P}(n-1)$	$\hat{\mathbf{x}}(n\|n) = \hat{\boldsymbol{\xi}}(\nu\|\nu-1) + \mathbf{K}(\nu)[\boldsymbol{\psi}(\nu) - \mathbf{X}(\nu)\hat{\boldsymbol{\xi}}(\nu\|\nu-1)]$
	$\mathbf{P}(n\|n) = [\mathbf{I} - \mathbf{K}(\nu)\mathbf{X}(\nu)]\mathbf{\Pi}(\nu\|\nu-1)$

State-Space Formulation of the RLS Problem

We begin by formulating the state-space description for the underlying dynamics of RLS filters. Consider the special case of an RLS filter with the forgetting factor, $\lambda = 1$. Recall in our convergence analysis of RLS, we had as our model

$$d(n) = \mathbf{w}_o^H\mathbf{u}(n) + e_o(n)$$

where $\mathbf{w}_0$ is the unknown optimal filter or parameter vector of the model and $e_o(n)$ is the associated modeling error or measurement error modeled as white noise. Note that the parameter vector is fixed.

With $\lambda = 1$, the transition matrix of the state-space model is equal to the identity matrix and since the underlying dynamics of the RLS filter are unforced, the process noise is zero. Using Kalman filter notation, we postulate the state-space model of RLS filters with $\lambda = 1$:

$$\mathbf{x}(n+1) = \mathbf{x}(n)$$
$$y(n) = \mathbf{C}(n)\mathbf{x}(n) + v(n)$$

where we assume the measurement noise, $v(n)$ is white with zero mean. A logical choice for $\mathbf{x}(n)$ (which is what we estimate with Kalman) is the parameter vector $\mathbf{w}_o$ (which is what RLS will want to adapt to). Comparing

$$d^*(n) = \mathbf{u}^H(n)\mathbf{w}_o + e_o^H(n)$$

to our measurement equation, we deduce the following identities for $\lambda = 1$:

$$\boxed{\begin{aligned}&\mathbf{x}(n)=\mathbf{w}_o\\&y(n)=d^*(n)\\&\mathbf{C}(n)=\mathbf{u}^H(n)\\&v(n)=e_o^*(n)\end{aligned}}.$$

With these identities, from the update algorithm for the Kalman filter (stochastic models), we can obtain the update algorithm for the RLS filter (deterministic models) as follows.

Since $\mathbf{A}(n)=\mathbf{I}$, we have

$$\hat{\mathbf{x}}(n|n-1)=\hat{\mathbf{x}}(n-1|n-1)$$

thus

$$\hat{\mathbf{x}}(n|n)=\hat{\mathbf{x}}(n-1|n-1)+\mathbf{K}(n)[\mathbf{y}(n)-\mathbf{C}(n)\hat{\mathbf{x}}(n-1|n-1)].$$

Since $\mathbf{A}(n)=\mathbf{I}$ and $w(n)$ is zero (unforced dynamics), we also have

$$\mathbf{P}(n|n-1)=\mathbf{P}(n-1|n-1)$$

thus

$$\mathbf{P}(n|n)=[\mathbf{I}-\mathbf{K}(n)\mathbf{C}(n)]\mathbf{P}(n-1|n-1).$$

Thus the first two equations in the Kalman filter are redundant and deleted.

The third equation is transformed as follows:

$$\begin{aligned}\mathbf{K}(n)&=\mathbf{P}(n|n-1)\mathbf{C}^H(n)[\mathbf{C}(n)\mathbf{P}(n|n-1)\mathbf{C}^H(n)+\mathbf{Q}_v(n)]^{-1}\\&=\mathbf{P}(n-1|n-1)\mathbf{C}^H(n)[\mathbf{C}(n)\mathbf{P}(n-1|n-1)\mathbf{C}^H(n)+\mathbf{Q}_v(n)]^{-1}\\&=\mathbf{P}(n-1)\mathbf{u}(n)[\mathbf{u}^H(n)\mathbf{P}(n-1)\mathbf{u}(n)+1]^{-1}\end{aligned}$$

or (why is $\mathbf{Q}_v(n)=1$—only specified zero mean and white?)

$$\boxed{\mathbf{k}(n)=\frac{\mathbf{P}(n-1)\mathbf{u}(n)}{1+\mathbf{u}^H(n)\mathbf{P}(n-1)\mathbf{u}(n)}}.$$

The fourth equation is transformed as follows:

$$\begin{aligned}\hat{\mathbf{x}}(n|n)&=\hat{\mathbf{x}}(n|n-1)+\mathbf{K}(n)[\mathbf{y}(n)-\mathbf{C}(n)\hat{\mathbf{x}}(n|n-1)]\\&=\hat{\mathbf{x}}(n-1|n-1)+\mathbf{K}(n)[\mathbf{y}(n)-\mathbf{u}^H(n)\hat{\mathbf{x}}(n-1|n-1)]\end{aligned}$$

or

$$\hat{\mathbf{w}}(n)=\hat{\mathbf{w}}(n-1)+\mathbf{k}(n)[d^*(n)-\mathbf{u}^H(n)\hat{\mathbf{w}}(n-1)].$$

This can be rewritten as

$$\boxed{\begin{aligned}&\xi(n)=d(n)-\hat{\mathbf{w}}^H(n-1)\mathbf{u}(n)\\&\hat{\mathbf{w}}(n)=\hat{\mathbf{w}}(n-1)+\mathbf{k}(n)\xi^*(n)\end{aligned}}.$$

The fifth equation is transformed as follows:

$$\mathbf{P}(n|n) = [\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]\mathbf{P}(n|n-1)$$
$$= [\mathbf{I} - \mathbf{K}(n)\mathbf{C}(n)]\mathbf{P}(n-1|n-1)$$

or

$$\boxed{\mathbf{P}(n) = [\mathbf{I} - \mathbf{k}(n)\mathbf{u}^H(n)]\mathbf{P}(n-1)}.$$

The above three boxed equations provide the update algorithm for the RLS filter.

A Comparison of Stochastic and Deterministic Models

The more general case of an RLS filter with $0 < \lambda \leq 1$ can also be solved in a similar way with a similar correspondence between Kalman and RLS variables

Table 13.2