Gradient for direct filter calibration

Take dynamics

$$X_{n+1} = AX_n + W_n$$

The observation/measurement is

$$Z_n = CX_n + V_n$$

The filter is

$$\widehat{X}_{n+1} = A\widehat{X}_n + K\left(Z_{n+1} - CA\widehat{X}_n\right)$$

The noise is Gaussian with mean zero and covariance

$$\mathrm{E}\left[\, W_n W_n^T \right] = R\ ,\ \ \mathrm{E}\left[\, V_n V_n^T \right] = S\ .$$

(Warning, S was called Q before. Here Q will be something different.) The estimation error is $Y_n = X_n - \widehat{X}_n$. Assume that $X_0 = 0$ and $\widehat{X}_0 = 0$. Take K to be a fixed update matrix. Define the total random cost to be

$$F = \sum_{n=1}^{N} Y_n^T Y_n = \sum_{n=1}^{N} \|Y_n\|^2\ .$$

This is a random variable that depends on K and the random noise W_n and V_n. For K fixed, the expectation is

$$V_0(0, K) = \mathrm{E}\left[F\right]$$

The random cost starting at time j with $\widehat{X}_j = \widehat{x}$ is

$$F_j(\widehat{x}, K, V_{j+1}, \ldots, V_N, X_j, \ldots, X_N) = \sum_{n=j}^{N} Y_n^T Y_n$$

In this formula, $\widehat{X}_j = \widehat{x}$, so the $n = j$ term is $\|X_j - \widehat{x}\|^2$. In this definition, we assume the "initial" condition at time $n = j$ is given by x and $\widehat{x}$. The cost-to-go function is

$$V_j(\widehat{x}, K) = \mathrm{E}[\, F_j(\widehat{x}, K, \ldots)]\ .$$

The optimal filtering problem is to choose K to minimize $V_0(0, K)$.

We want the derivatives

$$\nabla_K V_0(0, K) = \frac{\partial}{\partial K} V_0(0, K)\ .$$

Here, K is a matrix, so the partial derivatives are with respect to the entries in K. If the measurement has m components and the state has n components, then K is an $n \times m$, and $\nabla_K V$ has nm components. Important quantities are the derivatives of the random cost and the cost-to-go functions with respect to x and $\widehat{x}$. We call these

$$P_j(\widehat{x},\ldots) = \nabla_{\widehat{x}} F_j(\widehat{x},\ldots)$$

and

$$p_j(\widehat{x}) = \nabla_{\widehat{x}} V_j(\widehat{x},\ldots) = \mathrm{E}[\, P_j(\widehat{x},K,\ldots)]$$

An algorithm for computing the derivative with respect to K uses P_j. The P_j are computed using a backward recurrence.

Start with

$$F_N = Y_N^T Y_N = (X_N - \widehat{x})^T \, (X_N - \widehat{x}) \ .$$

The derivative with respect to K is zero. The derivative with respect to $\widehat{x}$ (in numerator-layout) is

$$P_N = \nabla_{\widehat{x}} \, (X_N - \widehat{x})^T \, (X_N - \widehat{x}) = 2\,(\widehat{x} - X_N)^T$$

From this start, the rest of the derivatives are calculated using a backward recursion (back-propagation). For this calculation, use

$$F_j(\widehat{x},K,\ldots) = (X_j - \widehat{x})^T \, (X_j - \widehat{x}) + F_{j+1}(\widehat{X}_{j+1},\ldots) \ .$$

In this formula, $\widehat{X}_{j+1}$ is a function of K and $\widehat{x}$

$$\begin{aligned} \widehat{X}_{j+1} &= A\widehat{x} + K(Z_{j+1} - CA\widehat{x}) \\ &= (I - KC)A\widehat{x} + KZ_{j+1} \end{aligned}$$

The recurrence relation for P_j uses the chain rule and differentiates the second $\widehat{X}$ update formula

$$\nabla_{\widehat{x}} \widehat{X}_{j+1} = (I - KC)A$$

The chain rule gives (in numerator-layout)

$$\begin{aligned} P_j = \nabla_{\widehat{x}} F_j &= \nabla_{\widehat{x}} \, (X_j - \widehat{x})^T \, (X_j - \widehat{x}) + \nabla_{\widehat{x}} F_{j+1}(\widehat{X}_{j+1},\ldots) \\ &= 2\,(\widehat{x} - X_j)^T + \nabla_{\widehat{X}_{j+1}} F_{j+1}(\widehat{X}_{j+1},\ldots) \nabla_{\widehat{x}} \widehat{X}_{j+1} \\ &= 2\,(\widehat{x} - X_j)^T + P_{j+1}\,[(I - KC)A] \end{aligned}$$

The K derivative is similar. Define

$$Q_j = \nabla_K F_j(\widehat{x},K,\ldots)$$

Then in the numerator layout (using the first $\widehat{X}$ update formula)

$$\begin{aligned} Q_j &= \nabla_K \, (X_j - \widehat{x})^T \, (X_j - \widehat{x}) + \nabla_K F_{j+1}(\widehat{X}_{j+1},K,\ldots) + \nabla_{\widehat{X}_{j+1}} F_{j+1}(\widehat{X}_{j+1},K,\ldots) \nabla_K \widehat{X}_{j+1} \\ &= Q_{j+1} + P_{j+1} \cdot diag\{(Z_{j+1} - CA\widehat{x})\} \end{aligned}$$

In our Kalman filter example, P_j and Q_j will be row vectors if computed in this way. To get column vectors, simply take the transpose.