We now use the criterion of optimality to determine $K_k$. Since we will assume we know $p(y^o)$ and $p(x^i)$, we will choose a value for $K_k$ which minimizes the cost function $J$ (equation 12) for the minimum variance Bayes estimate. Specifically

$$J_k = E[(v^o_k)^T v^a_k]$$
$$= \text{trace}[P^a_k]$$

(19)

Evaluating $\partial J / \partial K_k = 0$ and solving for the so-called "Kalman Gain" matrix $K_k$ we have

$$K_k = P^a_k H_k^T [H_k P^a_k H_k^T + R_k]^{-1}$$

(20)

Substituting (20) into (18) then yields

$$P^a_k = [I - K_k H_k]P^i_k$$

(21)

Finally, using the state space equation (7)

$$x(t) = M(t, t_o)x(t_o) + \eta(t, t_o)$$

we then obtain the estimates of $x_k^o$ needed in (15) and $P_k^I$ needed in (21)

$$x_k^I = M_{k-1}x_{k-1}^a$$

(22)

$$P_k^I = M_{k-1}P_{k-1}^a M_{k-1}^T + Q_{k-1}$$

(23)

where $Q_{k-1} = E[\eta_{k-1}\eta_{k-1}^T]$, and $x_{k-1}^a$ and $P_{k-1}^a$ are the optimal outputs from the previous iteration of the filter. From our earlier discussion (Section 3), $Q$ could represent random forcing in the system model due to transport model errors.

To use the filter we must provide initial (a priori) estimates for $x$ and $P$. Then from any prior output estimates $(x_{k-1}^a, P_{k-1}^a)$, we use measurement $k$ information $(y_k^o, R_k)$ and model information $(H_k, Q_k)$ together with equations (22), (23), (20), (15), and (21) to provide outputs $x_k^o$ and $P_k^I$ for inputs to the next step. The filter equations are summarized in Table 1.

Some intuitive concepts regarding the DKF are useful in understanding its operation. First, from equation (20), the gain matrix $K_k \rightarrow H_k^{-1}$ (its "maximum" value) as the measurement error covariance (noise) matrix $R_k \rightarrow 0$, and $K_k \rightarrow P_k^I H_k^TR_k^{-1}$ (its "minimum" value) as $R_k \rightarrow \infty$. Since the update in the state vector $x_k^o - x_k^I$ varies linearly with $K_k$, it is clear that measurements noisy enough so that $R_k$ much exceeds $H_k P_k^I H_k^T$, will contribute much less to improvement of the state vector estimation.

In this respect we can usefully consider $H_k P_k^I H_k^T$ as the error covariance matrix for the measurement estimates $y_k$. This emphasizes the importance of the weighting
of the data inherent in \( R_k \) and the distortions created if erroneous \( R_k \) are used. Note that \( R_k \) can include model error, mismatch error, and instrumental error as noted earlier.

Second, using (21), and recognizing that the maximum value of \( K_k H_k = I \), we see \( P_k^i \leq P_k^j \) with equality occurring for infinitely noisy measurements. Hence, the error covariance matrix \( P_k \) (whose diagonal elements are the variances of the state vector element estimates) decreases by amounts sensitively dependent on the measurement errors.

Third, we note from (23), that random forcings \( \eta \) in the system (state-space) model [equation (7)], which are represented here by \( Q \), will increase the extrapolated error covariance matrix \( P_k^e \) by amounts depending on the relative values of \( Q_{k<1} \) and the extrapolation matrix \( M_{k-1} P_{k-1}^e M_{k-1}^T \) in the absence of system (state-space) model noise. The inclusion of \( Q \) lessens the influence (or memory) of previous iterations in the filter operation. In the extreme, sufficiently large values of \( Q \) will prevent the capability of even non-noisy measurements to decrease \( P_k \) and hence increase the confidence in the state vector estimate. In other words excellent (non-noisy) measurements are of little use if the system (state-space) model is very noisy (e.g., through random variations \( \eta \) introduced by random transport errors).
Table 1: Kalman Filter Equations

<table>
<thead>
<tr>
<th>Definition</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement equation (model)</td>
<td>$y_k^o = H_k x_k^a + \epsilon_k$; $y_k = H_k x_k^f$</td>
</tr>
<tr>
<td>System (state) equation (model)</td>
<td>$x_k = M_{k-1} x_{k-1} + \eta_{k-1}$</td>
</tr>
<tr>
<td>State update</td>
<td>$x_k^o - x_k^f = K_k (y_k^o - y_k)$</td>
</tr>
<tr>
<td>Error Update</td>
<td>$P_k^a = (1 - K_k H_k) P_k^f$</td>
</tr>
<tr>
<td>Kalman gain update</td>
<td>$K_k = P_k^f H_k^T (H_k P_k^f H_k^T + R_k)^{-1}$</td>
</tr>
<tr>
<td>State time extrapolation</td>
<td>$x_k^f = M_{k-1} x_{k-1}^a$</td>
</tr>
<tr>
<td>Error time extrapolation</td>
<td>$P_k^f = M_{k-1} P_{k-1}^a M_{k-1}^T + Q_{k-1}$</td>
</tr>
<tr>
<td>System random forcing covariance</td>
<td>$Q_k = E(\eta_k \eta_k^T)$</td>
</tr>
<tr>
<td>Measurement error covariance</td>
<td>$R_k = E(\epsilon_k \epsilon_k^T)$</td>
</tr>
<tr>
<td>Estimation error covariance</td>
<td>$P_k = E(\nu_k \nu_k^T)$</td>
</tr>
<tr>
<td>Input measurement matrix</td>
<td>$= H_k = \partial y_k / \partial x_k$</td>
</tr>
<tr>
<td>Input system random forcing</td>
<td>$= Q_k$</td>
</tr>
<tr>
<td>Input state extrapolation</td>
<td>$= M_k$</td>
</tr>
<tr>
<td>Input measurement</td>
<td>$= y_k^o$</td>
</tr>
<tr>
<td>Input measurement error covariance</td>
<td>$= R_k$</td>
</tr>
<tr>
<td>Filter iteration</td>
<td>$\rightarrow (k-1)^{\prime}, \rightarrow \text{estimate}$</td>
</tr>
<tr>
<td></td>
<td>$\rightarrow (k-1)^*; \rightarrow \text{extrapolate}$</td>
</tr>
<tr>
<td></td>
<td>$\rightarrow (k)^\prime, \rightarrow -$</td>
</tr>
</tbody>
</table>

*A superscript a or superscript f denotes respectively the value before (f) or after (a) an update of an estimate using measurements, and k denotes the measurement number. In general, errors are assumed random with zero mean and measurement and estimation errors are uncorrelated.*
Kalman Filter

Optimal Estimation

\[ \hat{x}_k^+ - \hat{x}_k^- = K_k [z_k - \hat{z}_k] \]

\[ K_k = \text{function}(R_k, P_k, H_k) \]

Notes on slide notation:

\( (+) = a, \ (-) = f, \ z = y, \ (\wedge) = \text{model} \)
Example

Estimation of lifetime of $\text{CH}_3\text{CCl}_3$ (and hence $\text{OH}$)

Principal sink is $\text{CH}_3\text{CCl}_3 + \text{OH} \rightarrow \text{CH}_2\text{CCl}_3 + \text{H}_2\text{O}$

\[ \tau_j = \text{chemical lifetime} = (k_j[H])^{-1} \]

\[ X_j = \text{state vector} = \frac{E_j}{[M_j]} - \frac{[\text{CH}_3\text{CCl}_3]_j}{\tau_j} - \text{ocean sink} \]

Ocean sink (effective lifetime of 85 years)
\[ \text{CH}_3\text{CCl}_3 (g) \rightleftharpoons \text{CH}_3\text{CCl}_3 (aq) \rightarrow \text{hydrolysis products} \]

Figure by MIT OCW.

Image removed due to copyright considerations.

Method:
(1) Multiply reference set of $\frac{1}{C}$ by unknown parameter $\lambda$
(2) Multiply mole fractions by unknown calibration factor $\gamma$
(3) Estimate state vector $\vec{X} = (\vec{x})$
(4) Compute global average $\bar{C}$ from $\alpha$

Notes:
(1) $P_k(\cdot)$ is a probability distribution
(2) $R_k$ is the error in the measurement
(3) $C$ is the concentration of the chemical
(4) $\Delta C$ is the change in concentration

Smoother combines the forward and backwards filters.
Checking your answer:
Compare $Y^0_k$ versus $Y_k$

Image removed due to copyright considerations.