UW COVID-19 Weather Map
Real time Forecasting Equations
Prof. Les Atlas
University of Washington
November 29, 2020

Notation

$\mathbf{z}_{\boldsymbol{k}}$ observation vector at time k (measurements)
$\mathbf{x}_{\boldsymbol{k}}$ True SEIR state vector at time k (true model states, which are not observable)
$\hat{\mathbf{x}}_k^-$ a priori (predicted) estimate of system state vector at time k, which comes from the SEIR model, just before time k.
$\mathbf{e}_k^- = \mathbf{x}_{\boldsymbol{k}} - \hat{\mathbf{x}}_k^-$ a priori estimation error, a measure of the accuracy of the SEIR model just before time k.
$\mathbf{P}_k^- = E[\mathbf{e}_k^- \mathbf{e}_k^{-T}]$ a priori estimation error covariance matrix just before time k.
$\hat{\mathbf{x}}_{\boldsymbol{k}}$ a posteriori estimate of system state vector at time k.
$\mathbf{e}_k = \mathbf{x}_{\boldsymbol{k}} - \hat{\mathbf{x}}_k$ a posteriori estimation error, a measure of the accuracy of the SEIR model at time k.
$\mathbf{P}_k = E[\mathbf{e}_k \mathbf{e}_k^T]$ a posteriori estimation error covariance matrix at time k.
$\mathbf{A}$ State transition matrix.
$\mathbf{H}$ Measurement (observation) matrix.
$\mathbf{w}_k$ SEIR model noise, assumed zero mean and hoped to be as small as possible.
$\mathbf{v}_k$ Measurement (observation) noise, assumed zero mean and hoped to be as small as possible.
$\mathbf{Q}$ Process noise covariance.
$\mathbf{R}$ Measurement noise covariance.
$\mathbf{K}_k$ Kalman gain matrix at time k.
$\boldsymbol{\nu}_k$ Innovation error at time k.
$\mathbf{S}_k = E[\boldsymbol{\nu}_k \boldsymbol{\nu}_k^T]$ Innovation covariance matrix at time k.

Institute for Health Metrics and Evaluation (IHME) Pandemic Compartment Model

The IHME compartment model is SEIR equations which are represented by a nonlinear discrete-time state space model of the process and the measurements

$$\hat{\mathbf{x}}_k^- = f(\hat{\boldsymbol{x}}_{k-1}, k-1) + \mathbf{w}_{k-1} = \begin{bmatrix} S_k^- \\ E_k^- \\ I_{1,k}^- \\ I_{2,k}^- \\ R_k^- \end{bmatrix} = \boldsymbol{f}\left(\begin{bmatrix} S_{k-1} \\ E_{k-1} \\ I_{1,k-1} \\ I_{2,k-1} \\ R_{k-1} \end{bmatrix}\right) + \mathbf{w}_{k-1} \quad (1)$$

With the discrete notion for the SEIR equations we can represent the above $\boldsymbol{f}\left(\begin{bmatrix} S_{k-1} \\ E_{k-1} \\ I_{1,k-1} \\ I_{2,k-1} \\ R_{k-1} \end{bmatrix}\right)$ as the updates below.

$$
\begin{aligned}
S_k^- &= -\beta_{k-1}\frac{S_{k-1}\left(I_{1,k-1}+I_{2,k-1}\right)^{\alpha}}{N} \\
E_k^- &= \beta_{k-1}\frac{S_{k-1}\left(I_{1,k-1}+I_{2,k-1}\right)^{\alpha}}{N} - \sigma E_{k-1} \\
I_{1,k}^- &= \sigma E_{k-1} - \gamma_1 I_{1,k-1} \\
I_{2,k}^- &= \gamma_1 I_{1,k-1} - \gamma_2 I_{2,k-1} \\
R_k^- &= \gamma_2 I_{2,k-1}
\end{aligned}
$$

$0 < \alpha \leq 1$ is a predetermined mixing coefficient.

Since $\boldsymbol{f}\left(\begin{bmatrix} S_{k-1} \\ E_{k-1} \\ I_{1,k-1} \\ I_{2,k-1} \\ R_{k-1} \end{bmatrix}\right)$ is nonlinear we need to evaluate it's Jacobian's. This is the key change needed to make our approach follow what is typically called "An extended Kalman Filter." (Smith, Schmidt, & McGee, 1962) (McElhoe, 1966) The Jacobians are

$$
\boldsymbol{F}_{k-1} = \begin{bmatrix} \dfrac{\partial \boldsymbol{f}}{\partial S_{k-1}} & \dfrac{\partial \boldsymbol{f}}{\partial E_{k-1}} & \dfrac{\partial \boldsymbol{f}}{\partial I_{1,k-1}} & \dfrac{\partial \boldsymbol{f}}{\partial I_{2,k-1}} & \dfrac{\partial \boldsymbol{f}}{\partial R_{k-1}} \end{bmatrix}
$$

$$
= \begin{bmatrix}
-\beta_{k-1}\dfrac{\left(I_{1,k-1}+I_{2,k-1}\right)^{\alpha}}{N} & 0 & -\alpha\beta_{k-1}\dfrac{S_{k-1}\left(I_{1,k-1}+I_{2,k-1}\right)^{\alpha-1}}{N} & -\alpha\beta_{k-1}\dfrac{S_{k-1}\left(I_{1,k-1}+I_{2,k-1}\right)^{\alpha-1}}{N} & 0 \\
\beta_{k-1}\dfrac{\left(I_{1,k-1}+I_{2,k-1}\right)^{\alpha}}{N} & -\sigma & \alpha\beta_{k-1}\dfrac{S_{k-1}\left(I_{1,k-1}+I_{2,k-1}\right)^{\alpha-1}}{N} & \alpha\beta_{k-1}\dfrac{S_{k-1}\left(I_{1,k-1}+I_{2,k-1}\right)^{\alpha-1}}{N} & 0 \\
0 & \sigma & -\gamma_1 & 0 & 0 \\
0 & 0 & \gamma_1 & -\gamma_2 & 0 \\
0 & 0 & 0 & \gamma_2 & 0
\end{bmatrix}
$$

Updates via Daily Observations

The values we measure and forecast are represented by these measurement equations where the measured active case rate is to simply be proportional to the a priori estimate of the number of infected $I_{2,k}^-$ with proportionality constant ρ_2. And likewise, with a different proportion for the daily symptom survey, the number with symptoms is forecast to be $I_{1,k}^- = \rho_1 I_{2,k-1}$. A linear $\mathbf{H}$ is chosen so that $I_{1,k}^- = \rho_1 I_{2,k-1}$ and $I_{2,k}^- = \rho_1 I_{2,k-1}$.

In order to avoid matrix conditioning problems, the diagonal values of $\mathbf{H}$ were regularized a small amount instead using $\mathbf{H} + \varepsilon\mathbf{I}$ with small $\varepsilon > 0$ to yield the measurement equation

$$\mathbf{z}_k = \begin{bmatrix} 0 \\ 0 \\ z_{1,k} \\ z_{2,k} \\ 0 \end{bmatrix} = (\mathbf{H} + \varepsilon\mathbf{I})\hat{\mathbf{x}}_k^- + \mathbf{v}_k = \begin{bmatrix} \varepsilon & 0 & 0 & 0 & 0 \\ 0 & \varepsilon & 0 & 0 & 0 \\ 0 & 0 & \varepsilon & \rho_1 & 0 \\ 0 & 0 & 0 & \rho_2 & 0 \\ 0 & 0 & 0 & 0 & \varepsilon \end{bmatrix} \hat{\mathbf{x}}_k^- + \mathbf{v}_k \quad (2)$$

The compartment model error and measurement error random processes $\mathbf{w}_k$ and $\mathbf{v}_k$ are uncorrelated, zero-mean white-noise processes with assumed covariance matrices

$$\boldsymbol{E}[\mathbf{w}_k] = \mathbf{0}$$
$$\boldsymbol{E}[\mathbf{v}_k] = \mathbf{0}$$
$$\boldsymbol{E}[\mathbf{w}_k\mathbf{w}_l^T] = \begin{Bmatrix} \mathbf{Q}_k & k = l \\ 0 \text{ otherwise} \end{Bmatrix} = \boldsymbol{\delta}_{kl}\mathbf{Q}_k$$
$$\boldsymbol{E}[\mathbf{v}_k\mathbf{v}_l^T] = \begin{Bmatrix} \mathbf{R}_k & k = l \\ 0 \text{ otherwise} \end{Bmatrix} = \boldsymbol{\delta}_{kl}\mathbf{R}_k$$
$$\boldsymbol{E}[\mathbf{w}_k\mathbf{v}_l^T] = \mathbf{0} \text{ for all } k, l$$

where $\mathbf{Q}_k$ and $\mathbf{R}_k$ are symmetric positive semi-definite matrices.

This approach is based upon an integrated combination of the IHME SEIIR model and an extended Kalman filter. It combines the above equations via the usual predict and update cycles of the extended Kalman filtering algorithm.

Algorithm

Initialize, see page 18 of with sample-based estimates of

$$\hat{\mathbf{x}}_{k-1}^- = E[\hat{\boldsymbol{x}}_{k-1}], \mathbf{P}_{k-1} = E[(\mathbf{x}_{\boldsymbol{k-1}} - \hat{\mathbf{x}}_{k-1}^-)(\mathbf{x}_{\boldsymbol{k-1}} - \hat{\mathbf{x}}_{k-1}^-)^{\boldsymbol{T}}]$$

To estimate the above quantities, we can use the previous week's average case rates and model predictions $I_{2,k-1}^-$ to find $\hat{\mathbf{x}}_{k-1}^-$ and to estimate $\rho_2 = \frac{z_{2,k-1}}{I_{2,k-1}^-}$, and the last week's average symptom survey $I_{2,k-1}^-$ to find $\hat{\mathbf{x}}_{k-1}^-$ and to estimate $\rho_1 = \frac{z_{2,k-1}}{I_{2,k-1}^-}$. So past model $I_{2,k-1}^-$ can be used to estimate ρ_2 and ρ_1.

Toward an Update-Based Estimation Algorithm

We define, at time step k, an *a priori*[1] state vector estimate $\hat{\mathbf{x}}_k^-$, given results of process equation 1 prior to time step k. We then also define, at time step k, an *a posteriori* state vector estimate $\hat{\mathbf{x}}_{\boldsymbol{k}}$, given our only external measurement vector, $\mathbf{z}_{\boldsymbol{k}}$ from equation 2. *A priori* and *a posteriori* estimate errors can then be defined as

$$\mathbf{e}_k^- = \mathbf{x}_{\boldsymbol{k}} - \hat{\mathbf{x}}_k^-$$

and

$$\mathbf{e}_k = \mathbf{x}_{\boldsymbol{k}} - \hat{\mathbf{x}}_k \quad (3)$$

[1] The names *a priori* (prior) and *a posteriori* (posterior) come from Bayes theorem, as commonly used in DSP, machine learning, and modern statistics. To see more details, along with the related notion of likelihood, see, for example, http://blog.ivansiu.com/blog/2014/09/24/notes-on-maximum-likelihood/ .

Our county estimates get updated via these quantities, the *a priori* estimate error covariance matrix

$$\mathbf{P}_k^- = E[\mathbf{e}_k^- \mathbf{e}_k^{-T}]$$

And the *a posteriori* estimate error covariance matrix is

$$\mathbf{P}_k = E[\mathbf{e}_k \mathbf{e}_k^T] \quad (4)$$

where T signifies vector transpose.

In deriving the equations for the Kalman filter algorithm, an initial goal is to find an equation that computes an *a posteriori* state vector estimate $\hat{\mathbf{x}}_k$ as a linear combination of the *a priori* state vector estimate $\hat{\mathbf{x}}_k^-$ and the weighted difference between an actual external measurement vector, $\mathbf{z}_k$ and a measurement prediction $\mathbf{H}\hat{\mathbf{x}}_k^-$. This linear combination, with detail from Bayes rules, as seen in other more extensive references, is expressed as

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_k^-)) (5)$$

The difference $\mathbf{v}_k = (\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_k^-)$ in above equation 4 is called the measurement residual or innovation. A residual of zero means that our prediction is perfect.

The $n \times m$ matrix $\mathbf{K}_k$(not necessarily square) in above equation 5 is called the Kalman gain or blending factor. It is chosen to minimize the *a posteriori* error covariance $\mathbf{P}_k$ from equation 4. This minimization can be accomplished by first substituting equation 5 into the equation 3's definition for $\mathbf{e}_k$, substituting that into equation 4, doing the expectation, taking the derivative of the trace of the result with respect to gain matrix $\mathbf{K}_k$, setting that result to zero, and then solving for matrix $\mathbf{K}_k$. Details can be found within references cited in the technical report used for these Kalman filter notes,[2] on its page 3. One form of gain matrix $\mathbf{K}_k$that minimizes the *a posteriori* error covariance$\mathbf{P}_k$ from equation 4 is

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}^T(\mathbf{H}\mathbf{P}_k\mathbf{H}^T + \mathbf{R})^{-1} = \frac{\mathbf{P}_k^- \mathbf{H}^T}{\mathbf{H}\mathbf{P}_k\mathbf{H}^T + \mathbf{R}} \qquad (6)$$

Getting insight from equation 6 is important:
As the measurement error covariance vector $\mathbf{R}$ approaches zero, the gain matrix $\mathbf{K}_k$ weights the measurement residual more heavily. That is

$$\lim_{\mathbf{R}\to\mathbf{0}} \mathbf{K}_k = \mathbf{H}^{-1}.$$

Alternatively, as the *a priori* estimate error covariance $\mathbf{P}_k^-$ approaches zero, the Kalman gain weights the measurement residual less heavily. Namely

$$\lim_{\mathbf{P}_k^-\to\mathbf{0}} \mathbf{K}_k = \mathbf{0}.$$

The intuition about these two limiting cases suggests a tradeoff, *key to the function of the Kalman filter and the Kalman gain* $\mathbf{K}_k$:

1. As the measurement error covariance $\mathbf{R}$ approaches zero, the actual external measurement vector $\mathbf{z}_k$ is trusted more and more, yet the predicted measurement $\mathbf{H}\hat{\mathbf{x}}_k^-$ is trusted less and less.

[2] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME–Journal of Basic Engineering*, 82 (Series D): 35-45, 1960.

2. Alternatively, as the a priori estimate error covariance $\mathbf{P}_k^-$ approaches zero the actual measurement vector $\mathbf{z}_k$ is trusted less and less, while the predicted measurement $\mathbf{H}\hat{\mathbf{x}}_k^-$ is trusted more and more.

Summary of the Discrete Extended Kalman Filter Real Time Update Algorithm

The Kalman filter estimates a process by using a form of feedback control: the filter estimates the process state at some time and then obtains feedback in the form of (noisy) measurements. As such, the equations for the Kalman filter fall into two groups: time update equations and measurement update equations. The time update equations are responsible for projecting forward (in time) the current state and error covariance estimates to obtain the a priori estimates for the next time step. The measurement update equations are responsible for the feedback—i.e. for incorporating a new measurement into the *a priori* estimate to obtain an improved *a posteriori* estimate.

The time update equations can also be thought of as *predictor* equations, while the measurement update equations can be thought of as *corrector* equations. Indeed the final estimation algorithm resembles that of a *predictor-corrector* algorithm for solving numerical problems as shown below in Figure 1.

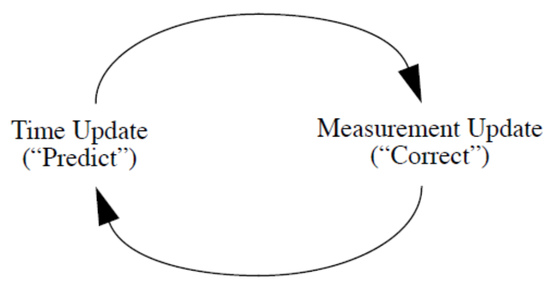

Figure 1. The ongoing discrete Kalman filter cycle. The *time update* projects the current state estimate ahead in time. The *measurement update* adjusts the projected estimate by an actual measurement at that time. (Figure from Welch and Bishop, "An Introduction to the Kalman Filter," TR 95-041, Department of Computer Science, University of North Carolina at Chapel Hill, 2006.)

The equations for the time and measurement updates are presented below.

Time update equations

Our SEIR Compartment model forecast updates are from the above equation 1:

$$\hat{\mathbf{x}}_k^- = f(\hat{\boldsymbol{x}}_{k-1}, k-1) = \begin{bmatrix} S_k^- \\ E_k^- \\ I_{1,k}^- \\ I_{2,k}^- \\ R_k^- \end{bmatrix} = \boldsymbol{f}\left(\begin{bmatrix} S_{k-1} \\ E_{k-1} \\ I_{1,k-1} \\ I_{2,k-1} \\ R_{k-1} \end{bmatrix}\right)$$

using the SEIR model discussed previously.

The estimation error covariance matrix time updates are

$$\mathbf{P}_k^- = \mathbf{F}_{k-1}\mathbf{P}_{k-1}\boldsymbol{F}_{\boldsymbol{k-1}}^{\boldsymbol{T}} + \mathbf{Q}$$

Where for the SEIR nonlinear compartment model we use its Jacobian $\mathbf{F}_{k-1} =$

$$\left[\frac{\partial f}{\partial S_{k-1}} \quad \frac{\partial f}{\partial E_{k-1}} \quad \frac{\partial f}{\partial I_{1,k-1}} \quad \frac{\partial f}{\partial I_{2,k-1}} \quad \frac{\partial f}{\partial R_{k-1}}\right] =$$

$$\begin{bmatrix} -\beta_{k-1}\frac{\left(I_{1,k-1}+I_{2,k-1}\right)^{\alpha}}{N} & 0 & -\alpha\beta_{k-1}\frac{S_{k-1}\left(I_{1,k-1}+I_{2,k-1}\right)^{\alpha-1}}{N} & -\alpha\beta_{k-1}\frac{S_{k-1}\left(I_{1,k-1}+I_{2,k-1}\right)^{\alpha-1}}{N} & 0 \\ \beta_{k-1}\frac{\left(I_{1,k-1}+I_{2,k-1}\right)^{\alpha}}{N} & -\sigma & \alpha\beta_{k-1}\frac{S_{k-1}\left(I_{1,k-1}+I_{2,k-1}\right)^{\alpha-1}}{N} & \alpha\beta_{k-1}\frac{S_{k-1}\left(I_{1,k-1}+I_{2,k-1}\right)^{\alpha-1}}{N} & 0 \\ 0 & \sigma & -\gamma_1 & 0 & 0 \\ 0 & 0 & \gamma_1 & -\gamma_2 & 0 \\ 0 & 0 & 0 & \gamma_2 & 0 \end{bmatrix}$$

The above two sets of equations project the state estimates $\hat{\mathbf{x}}_{k-1}$ and covariance estimates $\mathbf{P}_{k-1}$ forward from time step *k*-1 to time step *k*. The Jacobian matrix $\mathbf{F}_{k-1}$ is from above, and $\mathbf{Q}$ is from the previously defined, additive Gaussian noise covariance in the process model of equation 1, $p(\mathbf{w}_k) \sim N(\mathbf{0}, \mathbf{Q})$.

<u>Measurement update equations</u>

1. Kalman gain

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}^T (\mathbf{H}\mathbf{P}_k\mathbf{H}^T + \mathbf{R})^{-1} \text{ (7)}$$

2. Updated *a posteriori* state estimate

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_k^-) \text{ (8)}$$

3. Updated *a posteriori* covariance estimate

$$\mathbf{P}_k = (1 - \mathbf{K}_k\mathbf{H})\mathbf{P}_k^- \text{ (9)}$$

The 1st step in this measurement update is equation 7's computation of the Kalman gain $\mathbf{K}_k$, same as specified by equation 6. The next step, equation 8, finally uses the measured data vector $\mathbf{z}_k$ to generate the *a posteriori* state estimate $\hat{\mathbf{x}}_k$. Equation 8 is the same as previous equation 5. Above equation 9 then obtains the *a posteriori* error covariance matrix $\mathbf{P}_k$.

After each time and measurement update pair, the process is repeated with the previous *a posteriori* estimates used to project or predict the new a priori estimates. This recursive nature is one of the very appealing features of the Kalman filter—it makes practical implementations much more feasible than (for example) an implementation of a Wiener filter which is designed to operate on all of the data directly for each estimate. The Kalman filter instead recursively conditions the current estimate on all of the past measurements. Figure 2 below offers a complete picture of the operation of the filter, combining the above high-level diagram of Figure 1 with the update equations 7-9 from above.

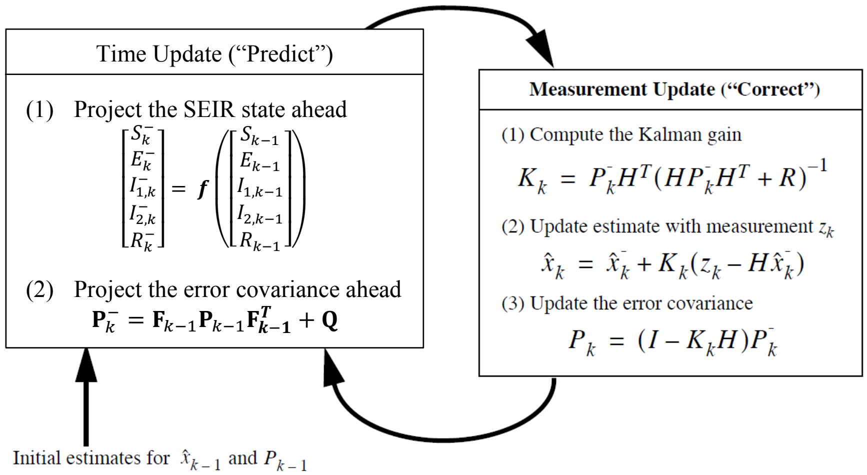

Figure 2. Starting from initial estimate, a complete picture of the operation of the Kalman filter.