Finding the Kalman gain matrix

- **KF step 2a**: Estimator (Kalman) gain matrix $L_k = \Sigma^-_{\tilde{x}\tilde{y},k}\Sigma^{-1}_{\tilde{y},k}$
  - To compute the Kalman gain, we must first compute several covariance matrices: we first find $\Sigma_{\tilde{y},k}$

$$\begin{aligned}\tilde{y}_k &= y_k - \hat{y}_k = C_k x_k + D_k u_k + v_k - C_k \hat{x}^-_k - D_k u_k = C_k \tilde{x}^-_k + v_k \\ \Sigma_{\tilde{y},k} &= \mathbb{E}\big[(C_k \tilde{x}^-_k + v_k)(C_k \tilde{x}^-_k + v_k)^T\big] \\ &= \mathbb{E}\big[C_k \tilde{x}^-_k (\tilde{x}^-_k)^T C_k^T + v_k (\tilde{x}^-_k)^T C_k^T + C_k \tilde{x}^-_k v_k^T + v_k v_k^T\big] \\ &= C_k \Sigma^-_{\tilde{x},k} C_k^T + \Sigma_{\tilde{v}}.\end{aligned}$$

  - Cross terms are zero since v_k is zero mean and is uncorrelated with $\tilde{x}^-_k$

Finding the Kalman gain matrix

- We now find $\Sigma^-_{\tilde{x}\tilde{y},k}$:

$$\begin{aligned}\mathbb{E}[\tilde{x}^-_k \tilde{y}^T_k] &= \mathbb{E}\big[\tilde{x}^-_k (C_k \tilde{x}^-_k + v_k)^T\big] = \mathbb{E}\big[\tilde{x}^-_k (\tilde{x}^-_k)^T C_k^T + \tilde{x}^-_k v_k^T\big] \\ &= \Sigma^-_{\tilde{x},k} C_k^T\end{aligned}$$

- Combining $\Sigma_{\tilde{y},k}$ and $\Sigma^-_{\tilde{x}\tilde{y},k}$ in $L_k = \Sigma^-_{\tilde{x}\tilde{y},k}\Sigma^{-1}_{\tilde{y},k}$,

$$L_k = \Sigma^-_{\tilde{x},k} C_k^T [C_k \Sigma^-_{\tilde{x},k} C_k^T + \Sigma_{\tilde{v}}]^{-1}$$

- **INTUITION**: Note that the computation of L_k is the most critical aspect of Kalman filtering that distinguishes it from a number of other estimation methods
  - The whole reason for calculating covariance matrices is to be able to update L_k
  - L_k is time-varying: It adapts to give the best update to the state estimate based on present conditions

Finding the Kalman gain matrix

- **INTUITION** (continued): Recall that we use L_k in the equation $\hat{x}^+_k = \hat{x}^-_k + L_k(y_k - \hat{y}_k)$
  - The first component to L_k, $\Sigma^-_{\tilde{x}\tilde{y},k}$, indicates relative need for correction to $\hat{x}_k$ and how well individual states within $\hat{x}_k$ are coupled to the measurements
  - We see this clearly in $\Sigma^-_{\tilde{x}\tilde{y},k} = \Sigma^-_{\tilde{x},k} C_k^T$
  - $\Sigma^-_{\tilde{x},k}$ tells us about state uncertainty at the present time, which we hope to reduce as much as possible
    - A large entry in $\Sigma^-_{\tilde{x},k}$ means that the corresponding state is very uncertain and therefore would benefit from a large update
    - A small entry in $\Sigma^-_{\tilde{x},k}$ means that the corresponding state is very well known already and does not need as large an update

Finding the Kalman gain matrix

- **INTUITION** (continued): Continuing to look at $\Sigma^-_{\tilde{x}\tilde{y},k} = \Sigma^-_{\tilde{x},k} C_k^T$
  - The C_k^T term gives the coupling between state and output
    - Entries that are zero indicate that a particular state has no direct influence on a particular output and therefore an output prediction error should not directly update that state
    - Entries that are large indicate that a particular state is highly coupled to an output so has a large contribution to any measured output prediction error; therefore, that state would benefit from a large update

Finding the Kalman gain matrix

- **INTUITION** (continued): Still looking at $L_k = \Sigma^-_{\tilde{x}\tilde{y},k} \Sigma^{-1}_{\tilde{y},k}$,
  - $\Sigma_{\tilde{y}}$ tells us how certain we are that measurement is reliable
    - If $\Sigma_{\tilde{y}}$ is "large," we want small, slow updates
    - If $\Sigma_{\tilde{y}}$ is "small," we want big updates
    - This explains why we divide the Kalman gain matrix by $\Sigma_{\tilde{y}}$
  - The form of $\Sigma_{\tilde{y}} = [C_k \Sigma^-_{\tilde{x},k} C_k^T + \Sigma_{\tilde{v}}]$ can also be explained
    - $C_k \Sigma^-_{\tilde{x}} C_k^T$ indicates how error in state contributes to error in output estimate
    - $\Sigma_{\tilde{v}}$ term indicates uncertainty in sensor reading due to sensor noise
    - Since sensor noise is assumed independent of the state, uncertainty in $\tilde{y}_k = y_k - \hat{y}_k$ adds the uncertainty in y_k to the uncertainty in $\hat{y}_k$

Finding the Kalman gain matrix

- **KF step 2b**: State estimate measurement update
  - Computes *a posteriori* state estimate by updating *a priori* estimate using estimator gain and output prediction error $y_k - \hat{y}_k$

$$\hat{x}_k^+ = \hat{x}_k^- + L_k(y_k - \hat{y}_k)$$

- **INTUITION**: $\hat{y}_k$ is predicted measurement, based on present state prediction
  - Therefore, $y_k - \hat{y}_k$ is what is unexpected or new in the measurement
  - We call this term the innovation. The innovation can be due to a bad system model, state error, or sensor noise
  - So, we want to use this new information to update the state, but must be careful to weight it according to the value of the information it contains
  - L_k is the optimal blending factor, as we have already discussed

Finding the Kalman gain matrix

- **KF step 2c**: Error covariance measurement update
  - Finally, we update error covariance matrix

$$\begin{aligned}\Sigma^{+}_{\tilde{x},k} &= \Sigma^{-}_{\tilde{x},k} - L_k \Sigma_{\tilde{y},k} L_k^T = \Sigma^{-}_{\tilde{x},k} - L_k \Sigma_{\tilde{y},k} \Sigma^{-T}_{\tilde{y},k} \left(\Sigma^{-}_{\tilde{x}\tilde{y},k}\right)^T \\ &= \Sigma^{-}_{\tilde{x},k} - L_k C_k \Sigma^{-}_{\tilde{x},k} \\ &= (I - L_k C_k) \Sigma^{-}_{\tilde{x},k}\end{aligned}$$

- **INTUITION**: A covariance matrix is positive semi-definite, and $L_k \Sigma_{\tilde{y},k} L_k^T$ is also a positive-semi-definite form, and we are subtracting this from the predicted-state covariance matrix; therefore, the resulting covariance is "lower" than the pre-measurement covariance
  - Measurement update has decreased our uncertainty in state estimate

Summary

- Have now derived the entire Kalman filter algorithm
- Next week, you will learn how to implement and visualize KF
- **KEY POINT**: Repeating from before, recall that the estimator output comprises the state estimate $\hat{x}_k^+$ and error covariance estimate $\Sigma^{+}_{\tilde{x},k}$
  - That is, we have high confidence that the truth lies within $\hat{x}_k^+ \pm 3\sqrt{\mathrm{diag}(\Sigma^{+}_{\tilde{x},k})}$
- **COMMENT**: If a measurement is missed for some reason, then skip steps 2a–c for that iteration. That is, set $L_k = 0$ and $\hat{x}_k^+ = \hat{x}_k^-$ and $\Sigma^{+}_{\tilde{x},k} = \Sigma^{-}_{\tilde{x},k}$