Pathway to finding Kalman gain

- So far, we have derived state estimate $\hat{x}_k^+ = \hat{x}_k^- + \mathbb{E}\big[\tilde{x}_k^- \mid y_k\big]$
  - But, what is $\mathbb{E}\big[\tilde{x}_k^- \mid y_k\big]$?
- To evaluate this expression, we consider very generically the problem of finding $f(x \mid y)$ if x and y are jointly Gaussian vectors
- We combine x and y into an augmented vector X where

$$X = \begin{bmatrix} x \\ y \end{bmatrix}, \qquad \mathbb{E}\big[X\big] = \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}, \qquad \Sigma_{\widetilde{X}} = \begin{bmatrix} \Sigma_{\tilde{x}} & \Sigma_{\tilde{x}\tilde{y}} \\ \Sigma_{\tilde{y}\tilde{x}} & \Sigma_{\tilde{y}} \end{bmatrix}$$

- We will then apply this generic result for $f(x \mid y)$ to our specific problem of determining $f(\tilde{x}_k^- \mid y_k)$ by setting $x = \tilde{x}_k^-$ and $y = y_k$
- We can then use this conditional pdf to find $\mathbb{E}\big[\tilde{x}_k^- \mid y_k\big]$

Expanding the conditional pdf

- To proceed, we first write $f(x \mid y) = f(x, y)/f(y)$
- Note that this is proportional to

$$\frac{f(x,y)}{f(y)} \propto \frac{\exp\left(-\frac{1}{2}\left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}\right)^T \Sigma_{\widetilde{X}}^{-1}\left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}\right)\right)}{\exp\left(-\frac{1}{2}(y-\bar{y})^T \Sigma_{\tilde{y}}^{-1}(y-\bar{y})\right)},$$

  where the constant of proportionality is not important to subsequent operations
- Recall that $\exp(A)/\exp(B) = \exp(A - B)\ldots$ which we take advantage of next

Condensing the exponential term

- The combined exponential argument becomes

$$-\frac{1}{2}\left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}\right)^T \Sigma_{\widetilde{X}}^{-1}\left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}\right) + \frac{1}{2}(y-\bar{y})^T \Sigma_{\tilde{y}}^{-1}(y-\bar{y})$$

- To condense notation somewhat, define $\tilde{x} = x - \bar{x}$ and $\tilde{y} = y - \bar{y}$
- Then, the terms in the exponent become,

$$-\frac{1}{2}\begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix}^T \Sigma_{\widetilde{X}}^{-1}\begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} + \frac{1}{2}\tilde{y}^T \Sigma_{\tilde{y}}^{-1}\tilde{y}$$

Inverting covariance (step 1)

- To proceed, we must invert $\Sigma_{\widetilde{X}}$. To do so, we substitute the following transformation (which you can verify)

$$\begin{bmatrix} \Sigma_{\tilde{x}} & \Sigma_{\tilde{x}\tilde{y}} \\ \Sigma_{\tilde{y}\tilde{x}} & \Sigma_{\tilde{y}} \end{bmatrix} = \begin{bmatrix} I & \Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \Sigma_{\tilde{x}} - \Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}\Sigma_{\tilde{y}\tilde{x}} & 0 \\ 0 & \Sigma_{\tilde{y}} \end{bmatrix} \begin{bmatrix} I & 0 \\ \Sigma_{\tilde{y}}^{-1}\Sigma_{\tilde{y}\tilde{x}} & I \end{bmatrix}$$

- Then,

$$\begin{aligned} \begin{bmatrix} \Sigma_{\tilde{x}} & \Sigma_{\tilde{x}\tilde{y}} \\ \Sigma_{\tilde{y}\tilde{x}} & \Sigma_{\tilde{y}} \end{bmatrix}^{-1} &= \begin{bmatrix} I & 0 \\ \Sigma_{\tilde{y}}^{-1}\Sigma_{\tilde{y}\tilde{x}} & I \end{bmatrix}^{-1} \begin{bmatrix} \Sigma_{\tilde{x}} - \Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}\Sigma_{\tilde{y}\tilde{x}} & 0 \\ 0 & \Sigma_{\tilde{y}} \end{bmatrix}^{-1} \begin{bmatrix} I & \Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1} \\ 0 & I \end{bmatrix}^{-1} \\ &= \begin{bmatrix} I & 0 \\ -\Sigma_{\tilde{y}}^{-1}\Sigma_{\tilde{y}\tilde{x}} & I \end{bmatrix} \begin{bmatrix} (\Sigma_{\tilde{x}} - \Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}\Sigma_{\tilde{y}\tilde{x}})^{-1} & 0 \\ 0 & \Sigma_{\tilde{y}}^{-1} \end{bmatrix} \begin{bmatrix} I & -\Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1} \\ 0 & I \end{bmatrix} \end{aligned}$$

Inverting covariance (step 2)

- Defining Schur complement of Σ: $M = \Sigma_{\tilde{x}} - \Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}\Sigma_{\tilde{y}\tilde{x}}$,

$$\begin{aligned} \text{exponent} = -\frac{1}{2}\big(&\tilde{x}^T M^{-1}\tilde{x} - \tilde{x}^T M^{-1}\Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}\tilde{y} - \tilde{y}^T\Sigma_{\tilde{y}}^{-1}\Sigma_{\tilde{y}\tilde{x}}M^{-1}\tilde{x} \\ &+ \tilde{y}^T\Sigma_{\tilde{y}}^{-1}\Sigma_{\tilde{y}\tilde{x}}M^{-1}\Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}\tilde{y} + \tilde{y}^T\Sigma_{\tilde{y}}^{-1}\tilde{y} - \tilde{y}^T\Sigma_{\tilde{y}}^{-1}\tilde{y}\big) \end{aligned}$$

- Last two terms cancel; remaining terms can be grouped as

$$\text{exponent} = -\frac{1}{2}\big(\tilde{x} - \Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}\tilde{y}\big)^T M^{-1}\big(\tilde{x} - \Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}\tilde{y}\big)$$

- In terms of only the original variables, exponent is

$$-\frac{1}{2}\Big(x-\big(\bar{x}+\Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}(y-\bar{y})\big)\Big)^T\Big(\Sigma_{\tilde{x}}-\Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}\Sigma_{\tilde{y}\tilde{x}}\Big)^{-1}\Big(x-\big(\bar{x}+\Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}(y-\bar{y})\big)\Big)$$

A solution to the generic conditional Gaussian

- Comparing this exponent to Gaussian standard form

$$-\frac{1}{2}\Big(x-\big(\bar{x}+\Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}(y-\bar{y})\big)\Big)^T\Big(\Sigma_{\tilde{x}}-\Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}\Sigma_{\tilde{y}\tilde{x}}\Big)^{-1}\Big(x-\big(\bar{x}+\Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}(y-\bar{y})\big)\Big)$$

  - Infer that mean of $f(x \mid y)$ must be $\bar{x} + \Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}(y - \bar{y})$
  - Infer that covariance of $f(x \mid y)$ must be $\Sigma_{\tilde{x}} - \Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}\Sigma_{\tilde{y}\tilde{x}}$
- So, we conclude that

$$\begin{aligned} &f(x \mid y) \sim \mathcal{N}(\bar{x} + \Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}(y - \bar{y}), \Sigma_{\tilde{x}} - \Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}\Sigma_{\tilde{y}\tilde{x}}) \\ &\mathbb{E}[x \mid y] = \mathbb{E}[x] + \Sigma_{\tilde{x}\tilde{y}}\Sigma_{\tilde{y}}^{-1}\big(y - \mathbb{E}[y]\big) \end{aligned}$$

State update equation

- Applying this to our problem, when $y_k = \tilde{y}_k + \hat{y}_k$, we get

$$\begin{aligned}\mathbb{E}\big[\tilde{x}_k^- \mid y_k\big] &= \mathbb{E}\big[\tilde{x}_k^-\big] + \Sigma^-_{\tilde{x}\tilde{y},k}\Sigma^{-1}_{\tilde{y},k}\big(y_k - \mathbb{E}[y_k]\big)\\ &= \mathbb{E}\big[\tilde{x}_k^-\big] + \Sigma^-_{\tilde{x}\tilde{y},k}\Sigma^{-1}_{\tilde{y},k}\big(\tilde{y}_k + \hat{y}_k - \mathbb{E}[\tilde{y}_k + \hat{y}_k]\big)\\ &= 0 + \Sigma^-_{\tilde{x}\tilde{y},k}\Sigma^{-1}_{\tilde{y},k}\big(\tilde{y}_k + \hat{y}_k - (0 + \hat{y}_k)\big) = \underbrace{\Sigma^-_{\tilde{x}\tilde{y},k}\Sigma^{-1}_{\tilde{y},k}}_{L_k}\tilde{y}_k\end{aligned}$$

- Putting all of the pieces together, we get the general update equation:

$$\hat{x}_k^+ = \hat{x}_k^- + L_k\tilde{y}_k$$

Uncertainty of state estimate

- State-estimate covariance (uncertainty) $\Sigma^+_{\tilde{x},k}$ is

$$\begin{aligned}\Sigma^+_{\tilde{x},k} &= \mathbb{E}\big[(x_k - \hat{x}_k^+)(x_k - \hat{x}_k^+)^T\big] = \mathbb{E}\Big[\big\{(x_k - \hat{x}_k^-) - L_k\tilde{y}_k\big\}\big\{(x_k - \hat{x}_k^-) - L_k\tilde{y}_k\big\}^T\Big]\\ &= \mathbb{E}\Big[\big\{\tilde{x}_k^- - L_k\tilde{y}_k\big\}\big\{\tilde{x}_k^- - L_k\tilde{y}_k\big\}^T\Big]\\ &= \Sigma^-_{\tilde{x},k} - L_k\underbrace{\mathbb{E}\big[\tilde{y}_k(\tilde{x}_k^-)^T\big]}_{\Sigma_{\tilde{y},k}L_k^T} - \underbrace{\mathbb{E}\big[\tilde{x}_k^-\tilde{y}_k^T\big]}_{L_k\Sigma_{\tilde{y},k}}L_k^T + L_k\Sigma_{\tilde{y},k}L_k^T\\ &= \Sigma^-_{\tilde{x},k} - L_k\Sigma_{\tilde{y},k}L_k^T\end{aligned}$$

- Final output of Gaussian sequential probabilistic inference has two parts:
  1. <u>State estimate</u>. Best guess of present state value, $\hat{x}_k^+$
  2. <u>Covariance estimate</u>. Uncertainty of $\hat{x}_k^+$, $\Sigma^+_{\tilde{x},k}$ can yield error bounds

Summary

- Generic Gaussian sequential probabilistic inference recursion:

$$\begin{aligned}\hat{x}_k^+ &= \hat{x}_k^- + L_k\big(y_k - \hat{y}_k\big) = \hat{x}_k^- + L_k\tilde{y}_k\\ \Sigma^+_{\tilde{x},k} &= \Sigma^-_{\tilde{x},k} - L_k\Sigma_{\tilde{y},k}L_k^T\end{aligned}$$

where

$$\begin{aligned}\hat{x}_k^- &= \mathbb{E}\big[x_k \mid \mathbb{Y}_{k-1}\big] & \Sigma^-_{\tilde{x},k} &= \mathbb{E}\big[(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)^T\big] = \mathbb{E}\big[(\tilde{x}_k^-)(\tilde{x}_k^-)^T\big]\\ \hat{x}_k^+ &= \mathbb{E}\big[x_k \mid \mathbb{Y}_k\big] & \Sigma^+_{\tilde{x},k} &= \mathbb{E}\big[(x_k - \hat{x}_k^+)(x_k - \hat{x}_k^+)^T\big] = \mathbb{E}\big[(\tilde{x}_k^+)(\tilde{x}_k^+)^T\big]\\ \hat{z}_k &= \mathbb{E}\big[z_k \mid \mathbb{Y}_{k-1}\big] & \Sigma_{\tilde{y},k} &= \mathbb{E}\big[(y_k - \hat{y}_k)(y_k - \hat{y}_k)^T\big] = \mathbb{E}\big[(\tilde{y}_k)(\tilde{y}_k)^T\big]\end{aligned}$$

$$L_k = \mathbb{E}\big[(x_k - \hat{x}_k^-)(y_k - \hat{y}_k)^T\big]\Sigma^{-1}_{\tilde{y},k} = \Sigma^-_{\tilde{x}\tilde{y},k}\Sigma^{-1}_{\tilde{y},k}$$

- Note that this is a linear recursion, even if the system is nonlinear(!)