16.346 Astrodynamics
Fall 2008

The Covariance Matrix #13.6

- Random Vector of Measurement Errors: $\boldsymbol{\alpha}$ where $\boldsymbol{\alpha}^{\mathrm{T}} = [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n]$
  Assume measurement errors independent.
- First and Second Moments: $E(\boldsymbol{\alpha}) = \overline{\boldsymbol{\alpha}} = \mathbf{0}$ and $E(\alpha_i\alpha_j) = \overline{\alpha_i\alpha_j} = 0 \quad (i \neq j)$
- Variances: $E(\alpha_i^2) = \overline{\alpha_i^2} = \sigma_i^2$
- Variance Matrix: $E(\boldsymbol{\alpha}\,\boldsymbol{\alpha}^{\mathrm{T}}) = \overline{\boldsymbol{\alpha}\,\boldsymbol{\alpha}^{\mathrm{T}}} = \mathbf{A} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}$
- Estimation Error Vector: $\boldsymbol{\epsilon} = \mathbf{PHA}^{-1}\boldsymbol{\alpha}$
- Covariance Matrix of Estimation Errors: $\boxed{E(\boldsymbol{\epsilon}\,\boldsymbol{\epsilon}^{\mathrm{T}}) = \overline{\boldsymbol{\epsilon}\,\boldsymbol{\epsilon}^{\mathrm{T}}}}$

$$\begin{aligned}
\boldsymbol{\epsilon}\boldsymbol{\epsilon}^{\mathrm{T}} &= \mathbf{PHA}^{-1}\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{P} \\
\overline{\boldsymbol{\epsilon}\boldsymbol{\epsilon}^{\mathrm{T}}} &= \mathbf{PHA}^{-1}\,\overline{\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}}}\,\mathbf{A}^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{P} = \mathbf{PHA}^{-1}\mathbf{AA}^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{P} \\
&= \mathbf{PHA}^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{P} = \mathbf{PP}^{-1}\mathbf{P} = \mathbf{P} = (\mathbf{P}^{-1})^{-1}
\end{aligned}$$

$$\boxed{\textbf{Covariance Matrix} = (\textbf{Information Matrix})^{-1}}$$

A Matrix Identity (The Magic Lemma)

Let $\mathbf{X}_{mn}$ and $\mathbf{Y}_{nm}$ be rectangular compatible matrices such that $\mathbf{X}_{mn}\mathbf{Y}_{nm}$ and $\mathbf{Y}_{nm}\mathbf{X}_{mn}$ are both meaningful.

However, $\mathbf{R}_{mm} = \mathbf{X}_{mn}\mathbf{Y}_{nm}$ is an $m \times m$ matrix while $\mathbf{S}_{nn} = \mathbf{Y}_{nm}\mathbf{X}_{mn}$ is an $n \times n$ matrix. With this understanding, the following sequence of matrix operations leads to a remarkable and very useful identity:

$$\begin{aligned}
(\mathbf{I}_{mm} + \mathbf{X}_{mn}\mathbf{Y}_{nm})(\mathbf{I}_{mm} + \mathbf{X}_{mn}\mathbf{Y}_{nm})^{-1} &= \mathbf{I}_{mm} \\
\mathbf{Y}_{nm}(\mathbf{I}_{mm} + \mathbf{X}_{mn}\mathbf{Y}_{nm})(\mathbf{I}_{mm} + \mathbf{X}_{mn}\mathbf{Y}_{nm})^{-1} &= \mathbf{Y}_{nm} \\
(\mathbf{I}_{nn} + \mathbf{Y}_{nm}\mathbf{X}_{mn})\mathbf{Y}_{nm}(\mathbf{I}_{mm} + \mathbf{X}_{mn}\mathbf{Y}_{nm})^{-1} &= \mathbf{Y}_{nm} \\
(\mathbf{I}_{nn} + \mathbf{Y}_{nm}\mathbf{X}_{mn})\mathbf{Y}_{nm}(\mathbf{I}_{mm} + \mathbf{X}_{mn}\mathbf{Y}_{nm})^{-1}\mathbf{X}_{mn} &= \mathbf{Y}_{nm}\mathbf{X}_{mn} \\
\mathbf{I}_{nn} + (\mathbf{I}_{nn} + \mathbf{Y}_{nm}\mathbf{X}_{mn})\mathbf{Y}_{nm}(\mathbf{I}_{mm} + \mathbf{X}_{mn}\mathbf{Y}_{nm})^{-1}\mathbf{X}_{mn} &= \mathbf{I}_{nn} + \mathbf{Y}_{nm}\mathbf{X}_{mn} \\
(\mathbf{I} + \mathbf{YX})^{-1}[\mathbf{I} + (\mathbf{I} + \mathbf{YX})\mathbf{Y}(\mathbf{I} + \mathbf{XY})^{-1}\mathbf{X}] &= (\mathbf{I} + \mathbf{YX})^{-1}(\mathbf{I} + \mathbf{YX}) \\
(\mathbf{I} + \mathbf{YX})^{-1} + \mathbf{Y}(\mathbf{I} + \mathbf{XY})^{-1}\mathbf{X} &= \mathbf{I}
\end{aligned}$$

Hence:

$$(\mathbf{I}_{nn} + \mathbf{Y}_{nm}\mathbf{X}_{mn})^{-1} = \mathbf{I}_{nn} - \mathbf{Y}_{nm}(\mathbf{I}_{mm} + \mathbf{X}_{mn}\mathbf{Y}_{nm})^{-1}\mathbf{X}_{mn}$$

To generalize: Let $\mathbf{Y}_{nm} = \mathbf{A}_{nn}\mathbf{B}_{nm}$ and $\mathbf{X}_{mn} = \mathbf{C}_{mm}^{-1}\mathbf{B}_{mn}^{\mathrm{T}}$. Then

$$\boxed{(\mathbf{A}^{-1} + \mathbf{BC}^{-1}\mathbf{B}^{\mathrm{T}})^{-1} = \mathbf{A} - \mathbf{AB}(\mathbf{C} + \mathbf{B}^{\mathrm{T}}\mathbf{AB})^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{A}}$$

Inverting the Information Matrix Using the Magic Lemma

- Recursive formulation: $\boxed{\mathbf{P}^{\star\,-1} = \mathbf{P}^{-1} + \mathbf{h}(\sigma^2)^{-1}\mathbf{h}^{\mathrm{T}}}$ $\qquad [= \mathbf{A}^{-1} + \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^{\mathrm{T}}]$

$$\boxed{(\mathbf{A}^{-1} + \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^{\mathrm{T}})^{-1} = \mathbf{A} - \mathbf{A}\mathbf{B}(\mathbf{C} + \mathbf{B}^{\mathrm{T}}\mathbf{A}\mathbf{B})^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{A}}$$

- Using the magic lemma: $\mathbf{P}^\star = \mathbf{P} - \mathbf{P}\mathbf{h}(\sigma^2 + \mathbf{h}^{\mathrm{T}}\mathbf{P}\mathbf{h})^{-1}\mathbf{h}^{\mathrm{T}}\mathbf{P}$
- Define: $a = \sigma^2 + \mathbf{h}^{\mathrm{T}}\mathbf{P}\mathbf{h}$ and $\mathbf{w} = \dfrac{1}{a}\mathbf{P}\mathbf{h}$ so that

$$\boxed{\mathbf{P}^\star = (\mathbf{I} - \mathbf{w}\mathbf{h}^{\mathrm{T}})\mathbf{P}}$$

The Square Root of the P Matrix #13.7

The matrix $\mathbf{W}$ is the **Square Root** of a Positive Definite Matrix $\mathbf{P}$ if $\mathbf{P} = \mathbf{W}\mathbf{W}^{\mathrm{T}}$

$$\begin{aligned}
\mathbf{P}^\star &= \mathbf{P} - \frac{1}{a}\mathbf{P}\mathbf{h}\mathbf{h}^{\mathrm{T}}\mathbf{P} \\
\mathbf{W}^\star\mathbf{W}^{\star\,\mathrm{T}} &= \mathbf{W}\left(\mathbf{I} - \frac{1}{a}\mathbf{W}^{\mathrm{T}}\mathbf{h}\mathbf{h}^{\mathrm{T}}\mathbf{W}\right)\mathbf{W}^{\mathrm{T}} \\
&= \mathbf{W}\left(\mathbf{I} - \frac{1}{a}\mathbf{z}\mathbf{z}^{\mathrm{T}}\right)\mathbf{W}^{\mathrm{T}} \\
&= \mathbf{W}(\mathbf{I} - \beta\mathbf{z}\mathbf{z}^{\mathrm{T}})(\mathbf{I} - \beta\mathbf{z}\mathbf{z}^{\mathrm{T}})\mathbf{W}^{\mathrm{T}} \\
&= \mathbf{W}(\mathbf{I} - 2\beta\mathbf{z}\mathbf{z}^{\mathrm{T}} + \beta^2\mathbf{z}\mathbf{z}^{\mathrm{T}}\mathbf{z}\mathbf{z}^{\mathrm{T}})\mathbf{W}^{\mathrm{T}} \\
&= \mathbf{W}[\mathbf{I} - (2\beta - \beta^2 z^2)\mathbf{z}\mathbf{z}^{\mathrm{T}}]\mathbf{W}^{\mathrm{T}}
\end{aligned}$$

But $\qquad z^2 = \mathbf{h}^{\mathrm{T}}\mathbf{W}\mathbf{W}^{\mathrm{T}}\mathbf{h} = \mathbf{h}^{\mathrm{T}}\mathbf{P}\mathbf{h} = a - \sigma^2$

Hence: $\qquad 2\beta - \beta^2 z^2 = \dfrac{1}{a} \quad \Longrightarrow \quad \beta = \dfrac{1}{a + \sqrt{a\sigma^2}}$

$$\boxed{\mathbf{W}^\star = \mathbf{W}\left(\mathbf{I} - \frac{\mathbf{z}\mathbf{z}^{\mathrm{T}}}{a + \sqrt{a\sigma^2}}\right) \qquad \mathbf{z} = \mathbf{W}^{\mathrm{T}}\mathbf{h}}$$

Properties of the Estimator

- Linear
- Unbiased: If measurements are exact ($\boldsymbol{\alpha} = \mathbf{0}$) then $\delta\widetilde{\mathbf{q}} = \delta\mathbf{q} = \mathbf{H}^{\mathrm{T}}\,\delta\mathbf{r}$ so that

$$\delta\widehat{\mathbf{r}} = \mathbf{P}\mathbf{P}^{-1}\,\delta\mathbf{r} = \delta\mathbf{r}$$

- Reduces to deterministic case ($\delta\widehat{\mathbf{r}} = \mathbf{H}^{-\mathrm{T}}\,\delta\widetilde{\mathbf{q}}$) if no redundant measurements.

  If $\mathbf{H}$ is square & non-singular, then $\mathbf{P} = \mathbf{H}^{-\mathrm{T}}\mathbf{A}\mathbf{H}^{-1}$ and

$$\delta\widehat{\mathbf{r}} = \mathbf{H}^{-\mathrm{T}}\mathbf{A}\mathbf{H}^{-1}\mathbf{H}\mathbf{A}^{-1}\,\delta\widetilde{\mathbf{q}} = \mathbf{H}^{-\mathrm{T}}\,\delta\widetilde{\mathbf{q}}$$

Recursive Form of the Estimator #13.6

Define $\widetilde{q}$

$$\delta\widehat{\mathbf{r}}^\star = \mathbf{F}^\star \delta\widetilde{\mathbf{q}}^\star \qquad \mathbf{F}^\star = \mathbf{P}^\star \mathbf{H}^\star \mathbf{A}^{\star\,-1} \qquad \delta\widetilde{\mathbf{q}}^\star = \begin{bmatrix} \delta\widetilde{\mathbf{q}} \\ \delta\widetilde{q} \end{bmatrix}$$

$$\mathbf{P}^\star = (\mathbf{I} - \mathbf{w}\mathbf{h}^\mathrm{T})\mathbf{P} \qquad \mathbf{H}^\star = \begin{bmatrix} \mathbf{H} & \mathbf{h} \end{bmatrix} \qquad \mathbf{A}^\star = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0}^\mathrm{T} & \sigma^2 \end{bmatrix}$$

Then

$$\mathbf{F}^\star = (\mathbf{I} - \mathbf{w}\mathbf{h}^\mathrm{T})\mathbf{P}\begin{bmatrix} \mathbf{H} & \mathbf{h} \end{bmatrix}\begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0}^\mathrm{T} & \sigma^{-2} \end{bmatrix} = (\mathbf{I} - \mathbf{w}\mathbf{h}^\mathrm{T})\mathbf{P}\begin{bmatrix} \mathbf{H}\mathbf{A}^{-1} & \dfrac{\mathbf{h}}{\sigma^2} \end{bmatrix}$$

$$= \begin{bmatrix} (\mathbf{I} - \mathbf{w}\mathbf{h}^\mathrm{T})\mathbf{F} & \dfrac{a\mathbf{w}}{\sigma^2} - \dfrac{\mathbf{w}(a-\sigma^2)}{\sigma^2} \end{bmatrix} = [\,(\mathbf{I} - \mathbf{w}\mathbf{h}^\mathrm{T})\mathbf{F} \quad \mathbf{w}\,]$$

$$\delta\widehat{\mathbf{r}}^\star = \mathbf{F}^\star\,\delta\widetilde{\mathbf{q}}^\star = \begin{bmatrix} (\mathbf{I} - \mathbf{w}\mathbf{h}^\mathrm{T})\mathbf{F} & \mathbf{w} \end{bmatrix}\begin{bmatrix} \delta\widetilde{\mathbf{q}} \\ \delta\widetilde{q} \end{bmatrix} = (\mathbf{I} - \mathbf{w}\mathbf{h}^\mathrm{T})\,\delta\widehat{\mathbf{r}} + \mathbf{w}\delta\widetilde{q}$$

$$= \delta\widehat{\mathbf{r}} + \mathbf{w}(\delta\widetilde{q} - \mathbf{h}^\mathrm{T}\,\delta\widehat{\mathbf{r}}) = \delta\widehat{\mathbf{r}} + \mathbf{w}(\delta\widetilde{q} - \delta\widehat{q})$$

Since $\delta q = \mathbf{h}^\mathrm{T}\,\delta\mathbf{r}$, then $\delta\widehat{q} = \mathbf{h}^\mathrm{T}\,\delta\widehat{\mathbf{r}}$ is the best estimate of the new measurement.

$$\begin{aligned} \delta\widehat{\mathbf{r}}^\star &= \delta\widehat{\mathbf{r}} + \mathbf{w}(\delta\widetilde{q} - \delta\widehat{q}) \\ \mathbf{w} &= \frac{1}{\sigma^2 + \mathbf{h}^\mathrm{T}\mathbf{P}\mathbf{h}}\,\mathbf{P}\mathbf{h} \\ \mathbf{P}^\star &= (\mathbf{I} - \mathbf{w}\mathbf{h}^\mathrm{T})\mathbf{P} \end{aligned}$$

or

$$\begin{aligned} \delta\widehat{\mathbf{r}}^\star &= \delta\widehat{\mathbf{r}} + \mathbf{w}(\delta\widetilde{q} - \delta\widehat{q}) \\ \mathbf{z} &= \mathbf{W}^\mathrm{T}\mathbf{h} \\ \mathbf{w} &= \frac{1}{\sigma^2 + z^2}\,\mathbf{W}\mathbf{z} \\ \mathbf{W}^\star &= \mathbf{W}\left(\mathbf{I} - \frac{\mathbf{z}\mathbf{z}^\mathrm{T}}{a + \sqrt{a\sigma^2}}\right) \end{aligned}$$

Triangular Square Root

$$\mathbf{W}\mathbf{W}^\mathrm{T} = \begin{bmatrix} 0 & 0 & w_1 \\ 0 & w_2 & w_3 \\ w_4 & w_5 & w_6 \end{bmatrix}\begin{bmatrix} 0 & 0 & w_4 \\ 0 & w_2 & w_5 \\ w_1 & w_3 & w_6 \end{bmatrix} = \begin{bmatrix} m_1 & m_2 & m_3 \\ m_2 & m_4 & m_5 \\ m_3 & m_5 & m_6 \end{bmatrix}$$

where

$$w_1^2 = m_1 \qquad w_3 = \frac{m_2}{w_1} \qquad w_6 = \frac{m_3}{w_1}$$

$$w_2^2 = \frac{m_1 m_4 - m_2^2}{m_1} \qquad w_4^2 = \frac{\det \mathbf{M}}{m_1 w_2^2} \qquad w_5 = \frac{m_5 - w_3 w_6}{w_2}$$