Examples of canonical info. (cont.)

(4) $y_4 = x_1 - x_2 + v_4 \qquad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

$$T_4 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \sigma^{-2} \qquad z_4 = \sigma^{-2} \begin{bmatrix} y_4 \\ -y_4 \end{bmatrix}$$

$$\bigoplus_{i=1}^{4} (T_i, z_i) = \left[\sigma^{-2} \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}, \ \sigma^{-2} \begin{bmatrix} y_1 + y_3 + y_4 \\ y_2 + y_3 - y_4 \end{bmatrix} \right]$$

case $(C) = (3) \oplus (4)$

(5) $y_5 = x + v_5 \qquad y_5 \in \mathbb{R}^2 \quad y_5 = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$

$$A = I \qquad S_5 = Var(v_5) = \sigma^2 \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}$$

$$T_5 = A_5^T S_5^{-1} A_5 = S_5^{-1} = \sigma^{-2} \frac{1}{1-r^2} \begin{bmatrix} 1 & -r \\ -r & 1 \end{bmatrix}$$

$$z_5 = A_5^T S_5^{-1} y_5 = \frac{\sigma^{-2}}{1-r^2} \begin{bmatrix} 1 & -r \\ -r & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

$$= \frac{1}{\sigma^2 (1-r^2)} \begin{bmatrix} z_1 - r z_2 \\ -r z_1 + z_2 \end{bmatrix}$$

(6) (case (e))

$$y_6 = \begin{bmatrix} x_1 + x_2 \\ x_1 - x_2 \end{bmatrix} + v_6 \qquad S_6 = \sigma^2 \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}$$

$$A_6 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad y_6 = \begin{bmatrix} z_3 \\ z_4 \end{bmatrix}$$

$$T_6 = A_6^T S_6^{-1} A = \frac{2}{\sigma^2} \begin{bmatrix} 1-r & 0 \\ 0 & 1+r \end{bmatrix}$$

(computed in (e))

$$z_6 = \underbrace{A_6^T S_6^{-1}}_{} y_6 = \frac{1}{\sigma^2} \begin{bmatrix} \frac{1}{1+r} & \frac{1}{1+r} \\ \frac{1}{1-r} & -\frac{1}{1-r} \end{bmatrix} \begin{bmatrix} z_3 \\ z_4 \end{bmatrix}$$

$$= \frac{1}{\sigma^2} \begin{bmatrix} \dfrac{z_3 + z_4}{1+r} \\ \dfrac{z_3 - z_4}{1-r} \end{bmatrix}$$

Eigen Basis.

Let $S: R \to R$ — operator

$Sx = \lambda x$ for some $x \in R$

$\lambda \in \mathbb{R}$

λ – eigenvalue for S

x – eigenvector

Th. If S – selfadjoint there exists an orthonormal basis of eigenvectors

$e_1, \dots, e_n$ $\quad Se_i = \lambda_i e_i$

$\lambda_1, \dots \lambda_n$

$\langle e_i, e_j \rangle = \delta_{ij}$

$$s_{ij} = \langle e_i, Se_j \rangle = \langle e_i, \lambda_j e_j \rangle$$

$$= \lambda_i \langle e_i, e_j \rangle = \lambda_i \delta_{ij} = \begin{cases} \lambda_i & i = j \\ 0 & i \neq j \end{cases}$$

$$\bar{S} = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} = \mathrm{diag}(\lambda_1 \dots \lambda_n)$$

$$\mathrm{tr}\, S = \sum_{i=1}^{n} \lambda_i$$

4

* Suppose $S \geq 0$ (non-neg definite / semi-pos def)
iff all $\lambda_i \geq 0$

$$\forall x \quad \langle Sx, x\rangle = \sum_{ij} S_{ij} x_j x_i$$

$$= \sum_i \lambda_i x_i^2 \geq 0 \ \ \forall x \text{ iff } \lambda_i \geq 0$$

* $S > 0$ (positive definite)
iff all $\lambda_i > 0$

$S>0$ means that $\forall x \neq 0 \ \langle Sx, x\rangle > 0$

* S is inventible iff
all $\lambda_i \neq 0$

$$S^{-1} = \begin{bmatrix} \lambda_1^{-1} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{-1} \end{bmatrix} \text{ - inverse of } \bar{S} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

* $\Rightarrow$ a positive-semidef operation
$S \geq 0$ is invertible iff it is $S > 0$

* S - self adjoint
  define $f(S)$ $f(x)$ - function of real x.

$$\overline{f(S)} = \begin{bmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) \end{bmatrix}$$

for example, $f(x) = \frac{1}{x} = x^{-1}$

$$\underline{f(S) = S^{-1}}$$

* $\overline{f(x) = \sqrt{x}}$ => def $S^{1/2}$

$$\overline{S^{1/2}} = \begin{bmatrix} \lambda_1^{1/2} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{1/2} \end{bmatrix}$$

$$S^{1/2} \cdot S^{1/2} = S$$

* if $S > 0$

$$S^{-1/2} = \left(S^{1/2}\right)^{-1} = \left(S^{-1}\right)^{1/2} > 0$$

$$S^{-1/2} \cdot S^{-1/2} = S^{-1}$$

If $S = \mathrm{Var}(v)$

$S \geq 0$

In the eigenbasis

$$\bar{S} = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix}$$

$$\bar{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

v_i - component in the eigenbasis.

Estimation with a priori information.

$$y = Ax + v \qquad v \sim (0, S)$$

a pioni knowledge about x:

$$x \sim (x_0, F) \qquad x = x_0 + \tilde{x} \quad \tilde{x} \sim (0, F)$$

$$\mathbf{E}_{pr} x = x_0 \quad Var_{pr}\, x = F \qquad \mathbf{E}_{pr} \tilde{x} = 0 \quad Var_{pr}\, \tilde{x} = F$$

$$Var\, x = F!\ \mathcal{D} \to \mathcal{D} \qquad x \in \mathcal{D}$$

$$y \in \mathcal{R}$$

Goal: Estimate x :

$$\hat{x} = Ry + r$$

need to find $R: \mathcal{R} \to \mathcal{D}$, $r \in \mathcal{D}$

$\hat{x}$ - as close to x as possible.

$$\begin{aligned}
\hat{x} - x &= Ry + r - x \\
&= R(Ax + v) + r - x \\
&= (RA - I)x + r + Rv
\end{aligned}$$

$$\mathbf{E}_v \|\hat{x} - x\|^2 = \mathbf{E}_v \|(RA-I)x + r + Rv\|^2$$

$$= \|(RA-I)x + r\|^2 + \mathbf{E}_v 2\langle (RA-I)x + r, Rv \rangle$$

$$+ \mathbf{E}_v \|Rv\|^2 = \|(RA-I)x + r\|^2 + tr\, RSR^*$$

$$H(R,r) = E_x\left(E_\nu \|\hat{x} - x\|^2\right) \qquad x = x_0 + \tilde{x}$$

$$= E_x \|(RA - I)x + r\|^2 + \operatorname{tr} RSR^*$$

$$= E_x \|(RA - I)(x_0 + \underline{\underline{\tilde{x}}}) + r\|^2 + \operatorname{tr}\ldots$$

$$= E_x \|(RA - I)\tilde{x} + [(RA - I)x_0 + r]\|^2 + \operatorname{tr}\ldots$$

$$= E_x \|(RA - I)\tilde{x}\|^2 +$$

$$+ \underline{E_x 2}\langle (RA - I)\underline{\underline{\tilde{x}}}, (RA - I)x_0 + r\rangle$$

$$+ \|(RA - I)x_0 + r\|^2 + \operatorname{tr} RSR^*$$

$$= \operatorname{tr} (RA - I)F(RA - I)^*$$

$$+ \|(RA - I)x_0 + \underline{\underline{r}}\|^2 + \operatorname{tr} RSR^*$$

$$\min_r H(R,r) = H(R) = \operatorname{tr}(RA - I)F(RA - I)^* + \operatorname{tr} RSR^*$$

$$r = (I - RA)x_0$$

$$H(R) = \operatorname{tr}\underbrace{\left[(RA - I)F(RA - I)^* + RSR^*\right]}_{= Q}$$

$$H(R) = \operatorname{tr} Q \sim \min_R .$$

$Q \sim \min_R$

$$Q = RAFA^*R^* - RAF - FA^*R^* + F + RSR^*$$

$$= \underbrace{R(AFA^* + S)R^*}_{\text{quadr wrt } R} - \underbrace{\underbrace{RAF} - \underbrace{FA^*}R^*}_{\text{linear}} + F$$

$$= RCR^* - RD^* - DR^* + F$$

$$C = AFA^* + S \qquad D = FA^*$$

$$Q = RCR^* - RD^* - DR^* + F$$

C - invertible.

S inv. $\Rightarrow$ pos-def $\overset{S>0}{\Rightarrow}$ $\underbrace{AFA^*}_{\geq 0} + \underbrace{S}_{>0} > 0$

$\Rightarrow C$ invertible

$$(R - DC^{-1})C(R - DC^{-1})^* =$$

$$= RCR^* - DC^{-1}CR^* - RCC^{-1}D^* + DC^{-1}CC^{-1}D^*$$

$$= \underbrace{RCR^* - DR^* - RD^*} + DC^{-1}D^*$$

$$Q = \underbrace{(R - DC^{-1})\, C\, (R - DC^{-1})^*}_{\geq 0} + F - DC^{-1}D^*$$

$\min Q$ when $\underline{R = DC^{-1}}$.

$$Q_{min} = F - DC^{-1}D^* \quad \Big| \quad \begin{array}{l} C = AFA^* + S \\ D = FA^* \end{array}$$

$$\left[\begin{array}{l} R = FA^*(AFA^*+S)^{-1} \\ Q_{min} = F - \underbrace{FA^*(AFA^*+S)^{-1}}AF \\ \qquad\ = F - RAF = (I - RA)F. \end{array}\right.$$

In Big Data context:

$\dim \mathcal{D} = m$ small.

$\dim \mathcal{R} = n$ large.

$AFA^* + S : \mathcal{R} \to \mathcal{R}$ $n \times n$ matrix — bad!

$$R = FA^*(AFA^* + S)^{-1}$$

$$= F^{1/2} \cdot F^{1/2} A^* \left(S^{1/2}\left(\underbrace{S^{-1/2} A F^{1/2}}_{C^*} \underbrace{F^{1/2} A^* S^{-1/2}}_{=C} + I\right) S^{1/2}\right)^{-1}$$

$$= F^{1/2} F^{1/2} A^* \left[S^{1/2}(\underbrace{C^*C}_{\geq 0} + \underbrace{I}_{>0}) S^{1/2}\right]^{-1}$$

$$= F^{1/2} \underbrace{F^{1/2} A^* S^{-1/2}}_{=C} (C^*C + I)^{-1} S^{-1/2}$$

$$= F^{1/2} C (C^*C + I)^{-1} S^{-1/2}$$

$$C(C^*C + I)^{-1} = (CC^* + I)^{-1} C$$

$C = S^{-1/2} A F^{1/2} : \mathcal{D} \to \mathcal{R}$

$C^*C + I \quad \mathcal{D} \to \mathcal{R}$

$CC^* + I : \mathcal{D} \to \mathcal{D}$

mult by $C^*C + I$ on the right

$CC^* + I$ on the left.

$$(CC^* + I)C = C(C^*C + I)$$

$$\underline{\underline{R}} = F^{1/2}(CC^* + I)^{-1} C S^{-1/2} \quad \Big| \quad C^* = S^{-1/2} A F^{1/2}, \quad C = F^{1/2} A^* S^{-1/2}$$

$$= F^{1/2}\Big(F^{1/2} A^* \underbrace{S^{-1/2} S^{-1/2}}_{S^{-1}} A F^{1/2} + I\Big)^{-1} F^{1/2} A^* \underbrace{S^{-1/2} S^{-1/2}}_{S^{-1}}$$

| Assume F invertible

$$= \Big(\underbrace{F^{-1/2} F^{1/2}} A^* S^{-1} A \underbrace{F^{1/2} F^{-1/2}} + F^{-1/2} F^{-1/2}\Big)^{-1} A^* S^{-1}$$

$$= \underbrace{\big(A^* S^{-1} A + F^{-1}\big)}\,^{-1} A^* S^{-1}$$

$$A^* S^{-1} A + F^{-1} : \mathcal{D} \to \mathcal{D} \quad m \times m$$

$$\underline{I - RA} = I - \big(A^* S^{-1} A + F^{-1}\big)^{-1} A^* S^{-1} A$$

$$= (\,)^{-1}(\,) - (\,)^{-1} A^* S^{-1} A$$

$$= (\,)^{-1}\big[\underbrace{A^* S^{-1} A} + F^{-1} \underbrace{- A^* S^{-1} A}\big]$$

$$= \big(A^* S^{-1} A + F^{-1}\big)^{-1} F^{-1}$$

$$\underline{\underline{r}} = (I - RA) x_0 = \big(A^* S^{-1} A + F^{-1}\big)^{-1} F^{-1} x_0$$

$$\underline{\underline{Q}} = (I - RA) F = \big(A^* S^{-1} A + F^{-1}\big)^{-1}$$

$$Q = \left(A^* S^{-1} A + F^{-1}\right)^{-1}$$

$$R = Q A^* S^{-1}$$

$$r = Q F^{-1} x_0$$

$$\hat{x} = Ry + r = Q\left(A^* S^{-1} y + F^{-1} x_0\right)$$

Vanishing a priori information

$F \to +\infty$ all eigenvalues $\to +\infty$

$\Rightarrow$ eigenvalues for $F^{-1} \to 0$

$\Rightarrow F^{-1} \to 0$

$$Q = \left(A^* S^{-1} A + \underbrace{F^{-1}}_{\to 0}\right)^{-1} \longrightarrow \left(A^* S^{-1} A\right)^{-1} = \mathrm{Var}(\hat{x}) \text{ for BLUE}$$

$$R = \left(A^* S^{-1} A + \underbrace{F^{-1}}_{\to 0}\right)^{-1} A^* S^{-1} = \left(A^* S^{-1} A\right)^{-1} A^* S^{-1}$$

$$r = \left(A^* S^{-1} A + \underbrace{F^{-1}}_{\to 0}\right) \underbrace{F^{-1}}_{\to 0} x_0 \to 0.$$

A priori information as an additional measurement.

$$x \sim (x_0, F)$$

$$x_0 = I \cdot x + \mu \qquad \mu \sim (0, F)$$

$$y = Ax + \nu \qquad \nu \sim (0, S)$$

$$\mathrm{Var}(\hat{x}) = (A^* S^{-1} A + I^* F^{-1} I)^{-1}$$
$$= (A^* S^{-1} A + F^{-1})^{-1} = \underline{\underline{Q}}$$

$$\hat{x} = \mathrm{Var}\,\hat{x}\,(A^* S^{-1} y + I^* F^{-1} x_0)$$
$$= (A^* S^{-1} A + F^{-1})^{-1} (A^* S^{-1} y + F^{-1} x_0)$$
$$= Ry + r$$

$$(y, A, S) \longmapsto (T, z) = (A^* S^{-1} A,\ A^* S^{-1} y)$$

$$(x_0, F) \longmapsto (T_0, z_0) = (F^{-1},\ F^{-1} x_0)$$

Transition from a priori to a posteriori info.

(x_0, F_0) - a priori.

$$y_1 = A_1 x + v_1 \qquad v_1 \sim (0, S_1)$$

$(\bar{x}_1, \bar{F}_1)$ - a posteriori inf.

$$\bar{x}_1 = \hat{x} = (F_0^{-1} + A_1^* S_1^{-1} A_1)^{-1} (F_0^{-1} x_0 + A_1^* S_1^{-1} y_1)$$

$$\bar{F}_1 = Var(\hat{x}) = (F_0^{-1} + A_1^* S_1^{-1} A_1)^{-1}$$

$$y_2 = A_2 x + v_2 \qquad v_2 \sim (0, S_2)$$

$$\bar{x}_2 = (F_0^{-1} + A_1^* S_1^{-1} A_1 + A_2^* S_2^{-1} A_2)^{-1} \times$$

$$\times (F_0^{-1} x_0 + A_1^* S_1^{-1} y_1 + A_2^* S_2^{-1} y_2)$$

$$= (\bar{F}_1^{-1} + A_2^* S_2^{-1} A_2)^{-1}$$

$$\cdot (\bar{F}_1^{-1} \bar{x}_1 + A_2^* S_2^{-1} y_2)$$

A priori $\to$ A posteriori info update.

$$
\begin{tikzcd}
 & (x_0, F_0) \arrow[d] & \\
(y_1, A_1, S_1) \arrow[r] & \oplus \arrow[r] & (\bar x_1, \bar F_1) \arrow[dl] \\
(y_2, A_2, S_2) \arrow[r] & \oplus \arrow[r] & (\bar x_2, \bar F_2) \\
 & \vdots & \\
 & & (\bar x_{k-1}, \bar F_{k-1}) \arrow[dl] \\
(y_k, A_k, S_k) \arrow[r] & \oplus \arrow[r] & (\bar x_k, \bar F_k)
\end{tikzcd}
$$

(x_0, F_0) – a priori

$(\bar x_1, \bar F_1)$ a post after mes. 1.

$$\bar F_k = \left(\bar F_{k-1}^{-1} + A_k^* S_k^{-1} A_k\right)^{-1}$$

$$\bar x_k = \bar F_k \left(\bar F_{k-1}^{-1} \bar x_{k-1} + A_k^* S_k^{-1} y_k\right)$$