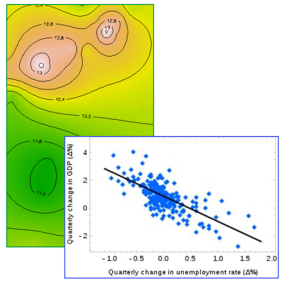
Least-Squares Estimation

Robert Stengel
Optimal Control and Estimation, MAE 546,
Princeton University, 2013

- Estimating unknown constants from redundant measurements
  - Least-squares
  - Weighted least-squares
- Recursive weighted least-squares estimator



Copyright 2013 by Robert Stengel. All rights reserved. For educational use only.
http://www.princeton.edu/~stengel/MAE3546.html
http://www.princeton.edu/~stengel/OptConEst.html

Perfect Measurement of a Constant Vector

- Given
  - Measurements, **y**, of a constant vector, **x**
- Estimate **x**
- Assume that output, **y**, is a perfect measurement and **H** is invertible

$$\mathbf{y} = \mathbf{H}\,\mathbf{x}$$

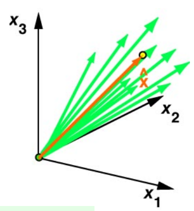
- **y**: (n x 1) output vector
- **H**: (n x n) output matrix
- **x** : (n x 1) vector to be estimated

- Estimate is based on inverse transformation

$$\hat{\mathbf{x}} = \mathbf{H}^{-1}\,\mathbf{y}$$

Imperfect Measurement of a Constant Vector

- Given
  - "Noisy" measurements, **z**, of a constant vector, **x**
- Effects of error can be reduced if measurement is redundant
- Noise-free output, **y**



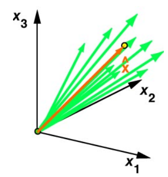
$$\mathbf{y} = \mathbf{H}\,\mathbf{x}$$

- **y**: (k x 1) output vector
- **H**: (k x n) output matrix, $k > n$
- **x** : (n x 1) vector to be estimated

- Measurement of output with error, **z**

$$\mathbf{z} = \mathbf{y} + \mathbf{n} = \mathbf{H}\,\mathbf{x} + \mathbf{n}$$

- **z**: (k x 1) measurement vector
- **n** : (k x 1) error vector



Cost Function for Least-Squares Estimate

- Measurement-error <u>residual</u>

$$\boldsymbol{\varepsilon} = \mathbf{z} - \mathbf{H}\,\hat{\mathbf{x}} = \mathbf{z} - \hat{\mathbf{y}} \qquad \dim(\boldsymbol{\varepsilon}) = (k \times 1)$$

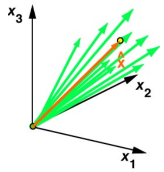
- Squared measurement error = cost function, J

Quadratic norm

$$J = \frac{1}{2}\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon} = \frac{1}{2}\left(\mathbf{z} - \mathbf{H}\,\hat{\mathbf{x}}\right)^T\left(\mathbf{z} - \mathbf{H}\,\hat{\mathbf{x}}\right)$$
$$= \frac{1}{2}\left(\mathbf{z}^T\mathbf{z} - \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{z} - \mathbf{z}^T\mathbf{H}\,\hat{\mathbf{x}} + \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{H}\,\hat{\mathbf{x}}\right)$$

- What is the control parameter?

The estimate of **x** $\quad \boxed{\hat{\mathbf{x}}} \quad \dim(\hat{\mathbf{x}}) = (n \times 1)$



Static Minimization Provides Least-Squares Estimate

Error cost function

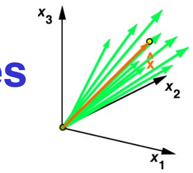
$$J = \frac{1}{2}\left(\mathbf{z}^T\mathbf{z} - \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{z} - \mathbf{z}^T\mathbf{H}\,\hat{\mathbf{x}} + \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{H}\,\hat{\mathbf{x}}\right)$$

Necessary condition

$$\frac{\partial J}{\partial \hat{\mathbf{x}}} = \mathbf{0} = \frac{1}{2}\left[\mathbf{0} - \left(\mathbf{H}^T\mathbf{z}\right)^T - \mathbf{z}^T\mathbf{H} + \left(\mathbf{H}^T\mathbf{H}\,\hat{\mathbf{x}}\right)^T + \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{H}\right]$$

$$\hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{H} = \mathbf{z}^T\mathbf{H}$$

Static Minimization Provides Least-Squares Estimate

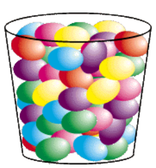


- Estimate is obtained using left pseudo-inverse matrix

$$\hat{\mathbf{x}}^T\left(\mathbf{H}^T\mathbf{H}\right)\left(\mathbf{H}^T\mathbf{H}\right)^{-1} = \hat{\mathbf{x}}^T = \mathbf{z}^T\mathbf{H}\left(\mathbf{H}^T\mathbf{H}\right)^{-1} \quad (row)$$

or

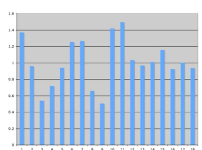
$$\hat{\mathbf{x}} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{z} \quad (column)$$



Example: Average Weight of a Pail of Jelly Beans

- Measurements are equally uncertain

$$z_i = x + n_i\ , \quad i = 1\ to\ k$$



- Express measurements as

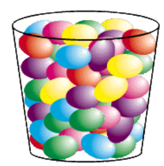
$$\mathbf{z} = \mathbf{H}x + \mathbf{n}$$

- Output matrix

$$\mathbf{H} = \begin{bmatrix} 1 \\ 1 \\ \ldots \\ 1 \end{bmatrix}$$

- Optimal estimate

$$\hat{x} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{z}$$



Average Weight of the Jelly Beans

Optimal estimate

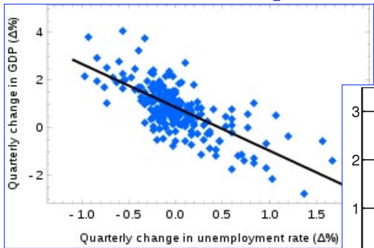
$$\hat{x} = \left(\begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ \ldots \\ 1 \end{bmatrix}\right)^{-1}\begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \\ \ldots \\ z_k \end{bmatrix}$$

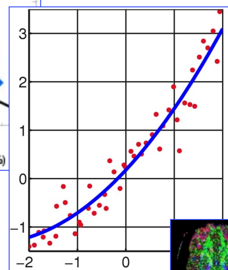
$$= (k)^{-1}\left(z_1 + z_2 + \ldots + z_k\right)$$

Simple average

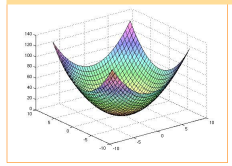
$$\hat{x} = \frac{1}{k}\sum_{i=1}^{k} z_i \quad [sample\ mean\ value]$$

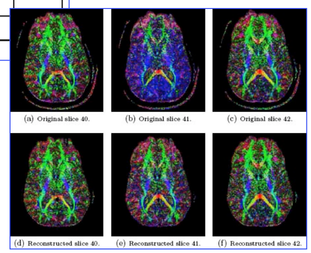
Least-Squares Applications





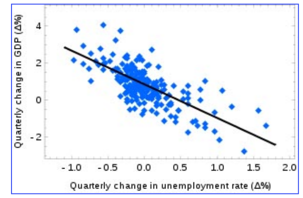
Error “Cost” Function





- More generally, least-squares estimation is used for
  - Higher-degree curve-fitting
  - Multivariate estimation

Least-Squares Linear Fit to Noisy Data



- Find trend line in noisy data

$$y = a_0 + a_1 x$$
$$z = (a_0 + a_1 x) + n$$

- Measurement vector

$$z_i = (a_0 + a_1 x) + n_i$$

$$\begin{bmatrix} z_1 \\ z_2 \\ \cdots \\ z_n \end{bmatrix} = \begin{bmatrix} (a_0 + a_1 x_1) \\ (a_0 + a_1 x_2) \\ \cdots \\ (a_0 + a_1 x_n) \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \cdots \\ n_n \end{bmatrix}$$

$$\mathbf{z} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \cdots & \cdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} + \mathbf{n} \triangleq \mathbf{Ha} + \mathbf{n}$$

- Error cost function

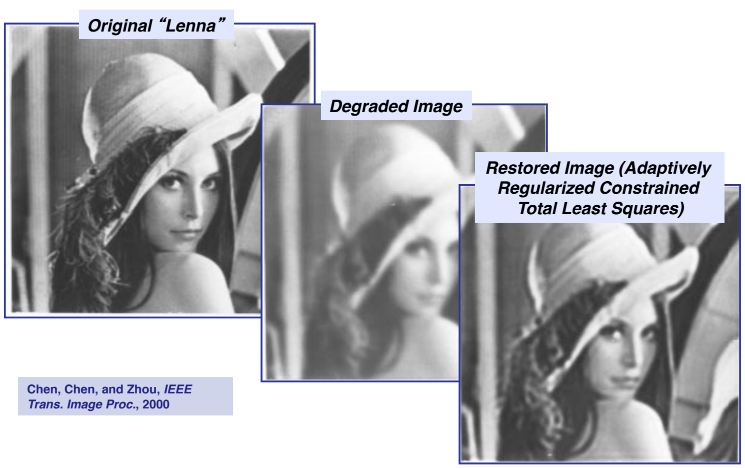
$$J = \frac{1}{2}(\mathbf{z} - \mathbf{H}\,\hat{\mathbf{a}})^T (\mathbf{z} - \mathbf{H}\,\hat{\mathbf{a}})$$

- Least-squares estimate of trend line
- Estimate ignores statistics of the error

$$\hat{\mathbf{a}} = \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \end{bmatrix} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{z}$$

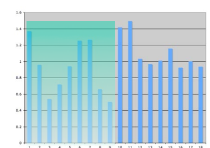
$$\hat{y} = \hat{a}_0 + \hat{a}_1 x$$

Least-Squares Image Processing



Chen, Chen, and Zhou, *IEEE Trans. Image Proc.*, 2000

Measurements of Differing Quality



- Suppose some elements of the measurement, **z**, are more uncertain than others

$$\mathbf{z} = \mathbf{Hx} + \mathbf{n}$$

- Give the more uncertain measurements less weight in arriving at the minimum-cost estimate
- Let **S** = measure of uncertainty; then express error cost in terms of $\mathbf{S}^{-1}$

$$J = \frac{1}{2}\boldsymbol{\varepsilon}^T\mathbf{S}^{-1}\boldsymbol{\varepsilon}$$

Error Cost and Necessary Condition for a Minimum

Error cost function, J

$$J = \frac{1}{2}\boldsymbol{\varepsilon}^T\mathbf{S}^{-1}\boldsymbol{\varepsilon} = \frac{1}{2}\left(\mathbf{z} - \mathbf{H}\,\hat{\mathbf{x}}\right)^T \mathbf{S}^{-1}\left(\mathbf{z} - \mathbf{H}\,\hat{\mathbf{x}}\right)$$
$$= \frac{1}{2}\left(\mathbf{z}^T\mathbf{S}^{-1}\mathbf{z} - \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{S}^{-1}\mathbf{z} - \mathbf{z}^T\mathbf{S}^{-1}\mathbf{H}\,\hat{\mathbf{x}} + \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{S}^{-1}\mathbf{H}\,\hat{\mathbf{x}}\right)$$

Necessary condition for a minimum

$$\frac{\partial J}{\partial \hat{\mathbf{x}}} = \mathbf{0}$$
$$= \frac{1}{2}\left[\mathbf{0} - \left(\mathbf{H}^T\mathbf{S}^{-1}\mathbf{z}\right)^T - \mathbf{z}^T\mathbf{S}^{-1}\mathbf{H} + \left(\mathbf{H}^T\mathbf{S}^{-1}\mathbf{H}\,\hat{\mathbf{x}}\right)^T + \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{S}^{-1}\mathbf{H}\right]$$

Weighted Least-Squares Estimate of a Constant Vector

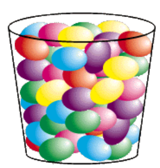
Necessary condition for a minimum

$$\left[\hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{S}^{-1}\mathbf{H} - \mathbf{z}^T\mathbf{S}^{-1}\mathbf{H}\right] = \mathbf{0}$$
$$\hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{S}^{-1}\mathbf{H} = \mathbf{z}^T\mathbf{S}^{-1}\mathbf{H}$$

Weighted left pseudo-inverse provides the solution

$$\hat{\mathbf{x}} = \left(\mathbf{H}^T\mathbf{S}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{S}^{-1}\mathbf{z}$$

The Return of the Jelly Beans



- Error-weighting matrix

$$\mathbf{S}^{-1} \triangleq \mathbf{A} = \begin{bmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & a_{kk} \end{bmatrix}$$

- Optimal estimate of average jelly bean weight

$$\hat{\mathbf{x}} = \left(\mathbf{H}^T\mathbf{S}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{S}^{-1}\mathbf{z}$$

$$\hat{x} = \left(\begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}\begin{bmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & a_{kk} \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ \ldots \\ 1 \end{bmatrix}\right)^{-1}\begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}\begin{bmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & a_{kk} \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \\ \ldots \\ z_k \end{bmatrix}$$

$$\hat{x} = \frac{\sum_{i=1}^{k} a_{ii} z_i}{\sum_{i=1}^{k} a_{ii}}$$

How to Chose the Error Weighting Matrix

a) Normalize the cost function according to expected measurement error, $\mathbf{S}_A$

$$J = \frac{1}{2}\boldsymbol{\varepsilon}^T\mathbf{S}_A^{-1}\boldsymbol{\varepsilon} = \frac{1}{2}\left(\mathbf{z} - \mathbf{y}\right)^T\mathbf{S}_A^{-1}\left(\mathbf{z} - \mathbf{y}\right) = \frac{1}{2}\left(\mathbf{z} - \mathbf{H}\,\mathbf{x}\right)^T\mathbf{S}_A^{-1}\left(\mathbf{z} - \mathbf{H}\,\mathbf{x}\right)$$

b) Normalize the cost function according to expected measurement residual, $\mathbf{S}_B$

$$J = \frac{1}{2}\boldsymbol{\varepsilon}^T\mathbf{S}_B^{-1}\boldsymbol{\varepsilon} = \frac{1}{2}\left(\mathbf{z} - \mathbf{H}\,\hat{\mathbf{x}}\right)^T\mathbf{S}_B^{-1}\left(\mathbf{z} - \mathbf{H}\,\hat{\mathbf{x}}\right)$$

Measurement Error Covariance, S_A

Expected value of <u>outer product</u> of measurement error vector

$$
\begin{aligned}
\mathbf{S}_A &= E\left[(\mathbf{z}-\mathbf{y})(\mathbf{z}-\mathbf{y})^T\right] \\
&= E\left[(\mathbf{z}-\mathbf{H}\mathbf{x})(\mathbf{z}-\mathbf{H}\mathbf{x})^T\right] \\
&= E\left[\mathbf{n}\mathbf{n}^T\right] \triangleq \mathbf{R}
\end{aligned}
$$

Measurement Residual Covariance, S_B

Expected value of <u>outer product</u> of measurement residual vector

$$
\begin{aligned}
\mathbf{S}_B &= E\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\right] \\
&= E\left[(\mathbf{z}-\mathbf{H}\hat{\mathbf{x}})(\mathbf{z}-\mathbf{H}\hat{\mathbf{x}})^T\right] \\
&= E\left[(\mathbf{H}\boldsymbol{\varepsilon}+\mathbf{n})(\mathbf{H}\boldsymbol{\varepsilon}+\mathbf{n})^T\right]
\end{aligned}
$$

$$\boldsymbol{\varepsilon} = (\mathbf{z}-\mathbf{H}\hat{\mathbf{x}})$$

$$
\begin{aligned}
&= \mathbf{H}E\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\right]\mathbf{H}^T + \mathbf{H}E\left(\boldsymbol{\varepsilon}\mathbf{n}^T\right) + E\left(\mathbf{n}\boldsymbol{\varepsilon}^T\right)\mathbf{H}^T + E\left(\mathbf{n}\mathbf{n}^T\right) \\
&\triangleq \mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{H}\mathbf{M} + \mathbf{M}^T\mathbf{H}^T + \mathbf{R}
\end{aligned}
$$

where

$$\mathbf{P} = E\left[(\mathbf{x}-\hat{\mathbf{x}})(\mathbf{x}-\hat{\mathbf{x}})^T\right]$$

$$\mathbf{M} = E\left[(\mathbf{x}-\hat{\mathbf{x}})\mathbf{n}^T\right]$$

$$\mathbf{R} = E\left[\mathbf{n}\mathbf{n}^T\right]$$

Requires iteration ("adaptation") of the estimate to find S_B

Recursive Least-Squares Estimation

- **Prior unweighted and weighted least-squares estimators use "batch-processing" approach**
  - All information is gathered prior to processing
  - All information is processed at once

- **Recursive approach**
  - Optimal estimate has been made from prior measurement set
  - New measurement set is obtained
  - Optimal estimate is improved by incremental change (or correction) to the prior optimal estimate

Prior Optimal Estimate

Initial measurement set and state estimate, with $S = S_A = R$

$$\mathbf{z}_1 = \mathbf{H}_1\mathbf{x} + \mathbf{n}_1$$

$$\hat{\mathbf{x}}_1 = \left(\mathbf{H}_1^T\mathbf{R}_1^{-1}\mathbf{H}_1\right)^{-1}\mathbf{H}_1^T\mathbf{R}_1^{-1}\mathbf{z}_1$$

$$\dim(\mathbf{z}_1) = \dim(\mathbf{n}_1) = k_1 \times 1$$

$$\dim(\mathbf{H}_1) = k_1 \times n$$

$$\dim(\mathbf{R}_1) = k_1 \times k_1$$

State estimate minimizes

$$J_1 = \frac{1}{2}\boldsymbol{\varepsilon}_1^T\mathbf{R}_1^{-1}\boldsymbol{\varepsilon}_1 = \frac{1}{2}\left(\mathbf{z}_1 - \mathbf{H}_1\hat{\mathbf{x}}_1\right)^T\mathbf{R}_1^{-1}\left(\mathbf{z}_1 - \mathbf{H}_1\hat{\mathbf{x}}_1\right)$$

New Measurement Set

New measurement

$$\mathbf{z}_2 = \mathbf{H}_2\mathbf{x} + \mathbf{n}_2$$

$\mathbf{R}_2$: Second measurement error covariance

Concatenation of old and new measurements

$$\mathbf{z} \triangleq \begin{pmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{pmatrix}$$

$$\dim(\mathbf{z}_2) = \dim(\mathbf{n}_2) = k_2 \times 1$$
$$\dim(\mathbf{H}_2) = k_2 \times n$$
$$\dim(\mathbf{R}_2) = k_2 \times k_2$$

Cost of Estimation Based on Both Measurement Sets

Cost function incorporates estimate made after incorporating $\mathbf{z}_2$

$$J_2 = \begin{bmatrix} (\mathbf{z}_1 - \mathbf{H}_1\hat{\mathbf{x}}_2)^T & (\mathbf{z}_2 - \mathbf{H}_2\hat{\mathbf{x}}_2)^T \end{bmatrix} \begin{pmatrix} \mathbf{R}_1^{-1} & 0 \\ 0 & \mathbf{R}_2^{-1} \end{pmatrix} \begin{bmatrix} (\mathbf{z}_1 - \mathbf{H}_1\hat{\mathbf{x}}_2) \\ (\mathbf{z}_2 - \mathbf{H}_2\hat{\mathbf{x}}_2) \end{bmatrix}$$
$$= (\mathbf{z}_1 - \mathbf{H}_1\hat{\mathbf{x}}_2)^T \mathbf{R}_1^{-1} (\mathbf{z}_1 - \mathbf{H}_1\hat{\mathbf{x}}_2) + (\mathbf{z}_2 - \mathbf{H}_2\hat{\mathbf{x}}_2)^T \mathbf{R}_2^{-1} (\mathbf{z}_2 - \mathbf{H}_2\hat{\mathbf{x}}_2)$$

Both residuals refer to $\hat{\mathbf{x}}_2$

Optimal Estimate Based on Both Measurement Sets

Simplification occurs because weighting matrix is block diagonal

$$\hat{\mathbf{x}}_2 = \left\{ \begin{bmatrix} \mathbf{H}_1^T & \mathbf{H}_2^T \end{bmatrix} \begin{bmatrix} \mathbf{R}_1^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_2^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{bmatrix} \right\}^{-1} \left\{ \begin{bmatrix} \mathbf{H}_1^T & \mathbf{H}_2^T \end{bmatrix} \begin{bmatrix} \mathbf{R}_1^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_2^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix} \right\}$$
$$= \left(\mathbf{H}_1^T\mathbf{R}_1^{-1}\mathbf{H}_1 + \mathbf{H}_2^T\mathbf{R}_2^{-1}\mathbf{H}_2\right)^{-1} \left(\mathbf{H}_1^T\mathbf{R}_1^{-1}\mathbf{z}_1 + \mathbf{H}_2^T\mathbf{R}_2^{-1}\mathbf{z}_2\right)$$

Apply Matrix Inversion Lemma

Define

$$\mathbf{P}_1^{-1} \triangleq \mathbf{H}_1^T\mathbf{R}_1^{-1}\mathbf{H}_1$$

Matrix inversion lemma

$$\left(\mathbf{H}_1^T\mathbf{R}_1^{-1}\mathbf{H}_1 + \mathbf{H}_2^T\mathbf{R}_2^{-1}\mathbf{H}_2\right)^{-1} =$$
$$\left(\mathbf{P}_1^{-1} + \mathbf{H}_2^T\mathbf{R}_2^{-1}\mathbf{H}_2\right)^{-1} =$$
$$\mathbf{P}_1 - \mathbf{P}_1\mathbf{H}_2^T\left(\mathbf{H}_2\mathbf{P}_1\mathbf{H}_2^T + \mathbf{R}_2\right)^{-1}\mathbf{H}_2\mathbf{P}_1$$

Improved Estimate Incorporating New Measurement Set

$$\hat{\mathbf{x}}_1 = \mathbf{P}_1\mathbf{H}_1^T\mathbf{R}_1^{-1}\mathbf{z}_1$$

New estimate is a correction to the old

$$\begin{aligned}\hat{\mathbf{x}}_2 &= \hat{\mathbf{x}}_1 - \mathbf{P}_1\mathbf{H}_2^T\left(\mathbf{H}_2\mathbf{P}_1\mathbf{H}_2^T + \mathbf{R}_2\right)^{-1}\mathbf{H}_2\hat{\mathbf{x}}_1 \\ &\quad + \mathbf{P}_1\mathbf{H}_2^T\left[\mathbf{I}_n - \left(\mathbf{H}_2\mathbf{P}_1\mathbf{H}_2^T + \mathbf{R}_2\right)^{-1}\mathbf{H}_2\mathbf{P}_1\mathbf{H}_2^T\right]\mathbf{R}_2^{-1}\mathbf{z}_2\end{aligned}$$

$$\begin{aligned}&= \left[\mathbf{I}_n - \left(\mathbf{H}_2\mathbf{P}_1\mathbf{H}_2^T + \mathbf{R}_2\right)^{-1}\mathbf{H}_2\mathbf{P}_1\mathbf{H}_2^T\right]\hat{\mathbf{x}}_1 \\ &+ \mathbf{P}_1\mathbf{H}_2^T\left[\mathbf{I}_n - \left(\mathbf{H}_2\mathbf{P}_1\mathbf{H}_2^T + \mathbf{R}_2\right)^{-1}\mathbf{H}_2\mathbf{P}_1\mathbf{H}_2^T\right]\mathbf{R}_2^{-1}\mathbf{z}_2\end{aligned}$$

Simplify Optimal Estimate Incorporating New Measurement Set

$$\mathbf{I} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1}, \quad \text{with} \quad \mathbf{A} \triangleq \mathbf{H}_2\mathbf{P}_1\mathbf{H}_2^T + \mathbf{R}_2$$

$$\begin{aligned}\hat{\mathbf{x}}_2 &= \hat{\mathbf{x}}_1 - \mathbf{P}_1\mathbf{H}_2^T\left(\mathbf{H}_2\mathbf{P}_1\mathbf{H}_2^T + \mathbf{R}_2\right)^{-1}\left(\mathbf{z}_2 - \mathbf{H}_2\hat{\mathbf{x}}_1\right) \\ &\triangleq \hat{\mathbf{x}}_1 - \mathbf{K}\left(\mathbf{z}_2 - \mathbf{H}_2\hat{\mathbf{x}}_1\right)\end{aligned}$$

$\mathbf{K}$: Estimator gain matrix

Recursive Optimal Estimate

- **Prior estimate may be based on prior incremental estimate, and so on**
- **Generalize to a recursive form, with sequential index *i***

$$\begin{aligned}\hat{\mathbf{x}}_i &= \hat{\mathbf{x}}_{i-1} - \mathbf{P}_{i-1}\mathbf{H}_i^T\left(\mathbf{H}_i\mathbf{P}_{i-1}\mathbf{H}_i^T + \mathbf{R}_i\right)^{-1}\left(\mathbf{z}_i - \mathbf{H}_i\hat{\mathbf{x}}_{i-1}\right) \\ &\triangleq \hat{\mathbf{x}}_{i-1} - \mathbf{K}_i\left(\mathbf{z}_i - \mathbf{H}_i\hat{\mathbf{x}}_{i-1}\right)\end{aligned}$$

with

$$\mathbf{P}_i = \left(\mathbf{P}_{i-1}^{-1} + \mathbf{H}_i^T\mathbf{R}_i^{-1}\mathbf{H}_i\right)^{-1}$$

$\dim(\mathbf{x}) = n \times 1;\quad \dim(\mathbf{P}) = n \times n$
$\dim(\mathbf{z}) = r \times 1;\quad \dim(\mathbf{R}) = r \times r$
$\dim(\mathbf{H}) = r \times n;\quad \dim(\mathbf{K}) = n \times r$

Example of Recursive Optimal Estimate

$$z = x + n$$

$$\hat{x}_i = \hat{x}_{i-1} + p_{i-1}\left(p_{i-1} + 1\right)^{-1}\left(z_i - \hat{x}_{i-1}\right)$$

$$k_i = p_{i-1}\left(p_{i-1} + 1\right)^{-1} = \frac{p_{i-1}}{\left(p_{i-1} + 1\right)}$$

$$p_i = \left(p_{i-1}^{-1} + 1\right)^{-1} = \frac{1}{\left({p_{i-1}}^{-1} + 1\right)}$$

$H = 1; \quad R = 1$

index	p-sub-i	k-sub-i
0	1	-
1	0.5	0.5
2	0.333	0.333
3	0.25	0.25
4	0.2	0.2
5	0.167	0.167

$$\begin{aligned}\hat{x}_0 &= z_0 \\ \hat{x}_1 &= 0.5\hat{x}_0 + 0.5z_1 \\ \hat{x}_2 &= 0.667\hat{x}_1 + 0.333z_2 \\ \hat{x}_3 &= 0.75\hat{x}_2 + 0.25z_3 \\ \hat{x}_4 &= 0.8\hat{x}_3 + 0.2z_4 \\ &\ldots\end{aligned}$$

Optimal Gain and Estimate-Error Covariance

$$\mathbf{P}_i = \left(\mathbf{P}_{i-1}^{-1} + \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{H}_i\right)^{-1}$$

$$\mathbf{K}_i = \mathbf{P}_{i-1} \mathbf{H}_i^T \left(\mathbf{H}_i \mathbf{P}_{i-1} \mathbf{H}_i^T + \mathbf{R}_i\right)^{-1}$$

- With constant estimation error matrix, **R**,
  - Error covariance decreases at each step
  - Estimator gain matrix, **K**, invariably goes to zero as number of samples increases
- Why?
- Each new sample has smaller effect on the average than the sample before

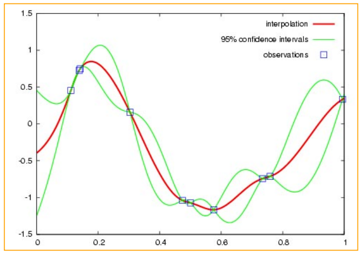
Next Time: Propagation of Uncertainty in Dynamic Systems

Supplemental Material

Weighted Least Squares ("Kriging") Estimates (Interpolation)

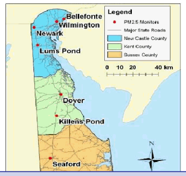
- Can be used with arbitrary interpolating functions

http://en.wikipedia.org/wiki/Kriging



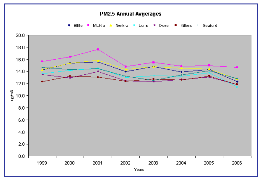
Weighted Least Squares Estimates of Particulate Concentration ($PM_{2.5}$)

Delaware Sampling Sites



Delaware Dept. of Natural Resources and Environmental Control, 2008

Delaware Average Concentrations



DE-NJ-PA $PM_{2.5}$ Estimates