Generalized Inverses & Least Squares Problems

B. Wayne Bequette

Two Related Problems

Problem 1

$$\min_x x^T x$$

$$s.t. Ax = b$$

Usually dim(x)>dim(b)
Similar to “coincidence pt. problem

Problem 2

$$\min_x ||Ax - b||$$

Usually dim(x)<dim(b)
Similar to “least squares” curve fit

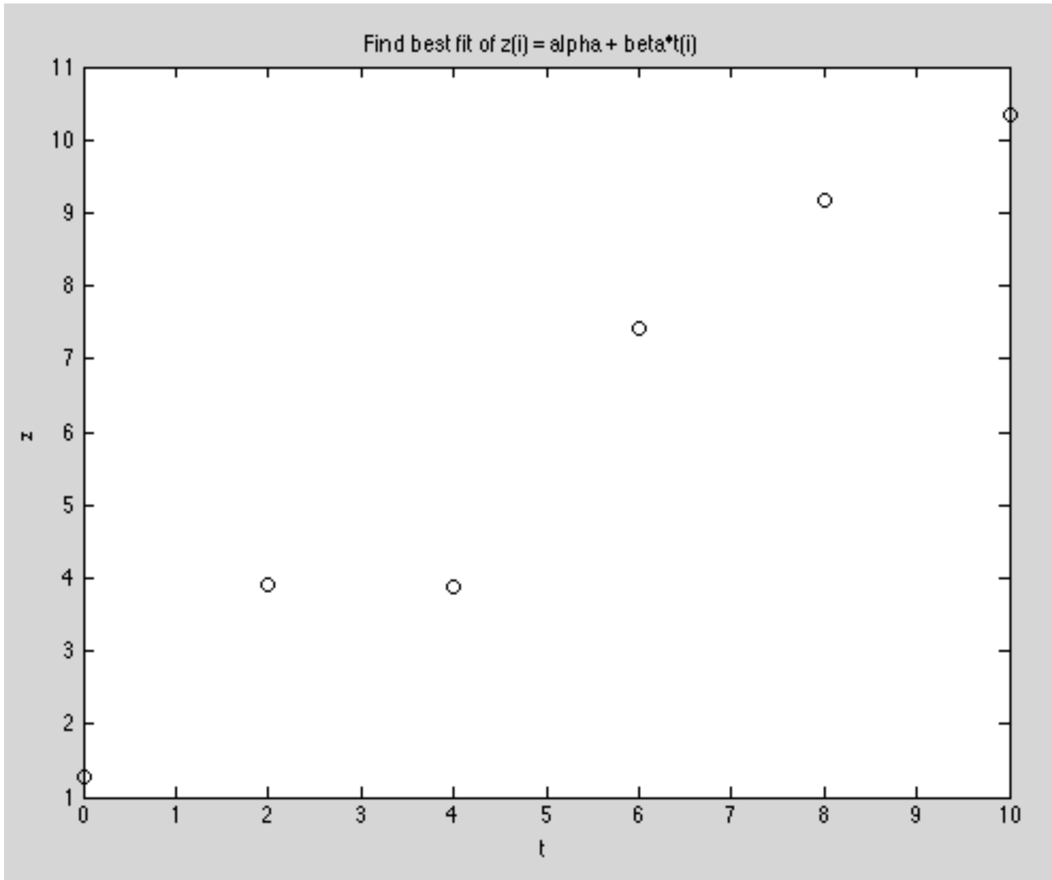
Same Solution

$$x = (A^T A)^{-1} A^T b$$

Not exactly true, as shown later:

$$x = A^T (AA^T)^{-1} b$$

Problem 2 example



$$\hat{z}_i = \propto + \beta t_i$$

$$\begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} z_1 \\ \vdots \\ z_m \end{bmatrix}$$

$$Ax = b$$

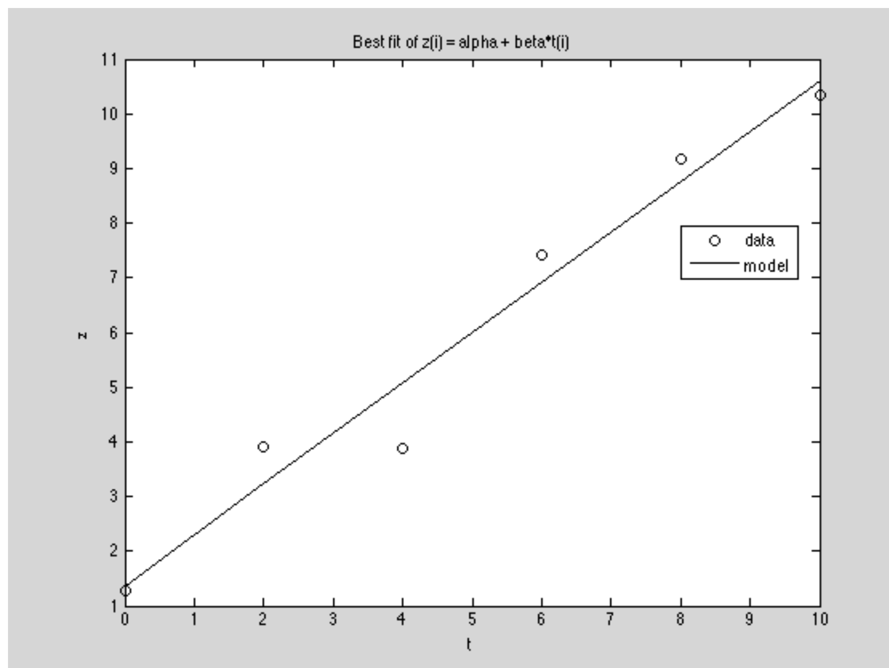
$$\begin{bmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 4 \\ 1 & 6 \\ 1 & 8 \\ 1 & 10 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1.27 \\ 3.92 \\ 3.87 \\ 7.43 \\ 9.16 \\ 10.35 \end{bmatrix}$$

$$x = (A^T A)^{-1} A^T b$$

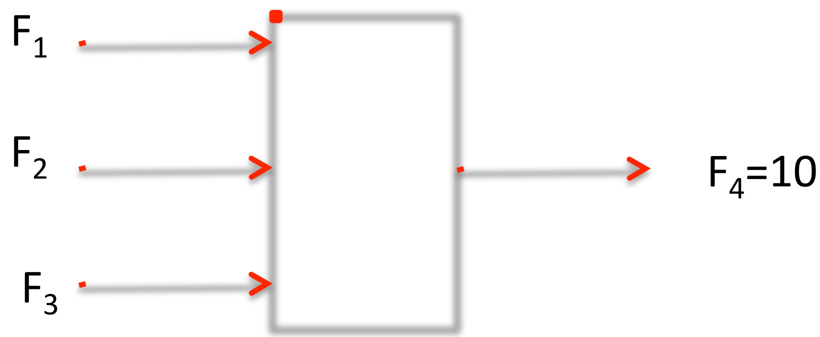
$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} & & \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_m \end{bmatrix}$$

(2x1) (2xm) (mx1)

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1.279 \\ 0.9239 \end{bmatrix}$$



Problem 1 Example



Objective Function

$$\min \sum_{i=1}^{3} F_i^2 \qquad \text{s.t.} \qquad F_1 + F_2 + F_3 = F_4$$

Material Balance

$$\min_x x^T x \qquad x^T x = [F_1 \quad F_2 \quad F_3]\begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} = \sum_{i=1}^{3} F_i^2$$

$$s.t.\, Ax = b \qquad [1 \quad 1 \quad 1]\begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} = [10]$$

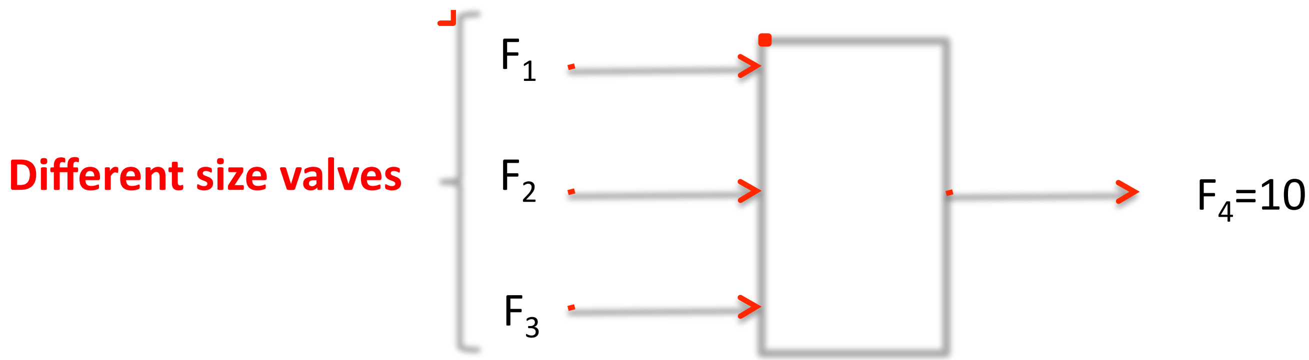
$$Ax = b$$

$$\underset{(3x1)}{x} = \underset{(3x1)(1x3)}{(A^T A)^{-1}} \underset{(3x1)(1x1)}{A^T b} \qquad A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Use SVD (MATLAB pinv)

$$\begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} = \begin{bmatrix} 3.333 \\ 3.333 \\ 3.333 \end{bmatrix}$$

Problem 1 Example – Weighted LS



Objective Function $\min_{F_1,F_2,F_3} \sum_{i=1}^{3} w_i F_i^2$ s.t. $F_1 + F_2 + F_3 = F_4$ **Material Balance**

$$\min_x x^T W x \qquad x^T W x = \begin{bmatrix} F_1 & F_2 & F_3 \end{bmatrix} \begin{bmatrix} w_1 & 0 & 0 \\ 0 & w_2 & 0 \\ 0 & 0 & w_3 \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} = \sum_{i=1}^{3} w_i F_i^2$$

$$s.t. Ax = b \qquad \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} = [10]$$

$Ax = b$

Weights $\begin{bmatrix} w_1 & 0 & 0 \\ 0 & w_2 & 0 \\ 0 & 0 & w_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 16 \end{bmatrix}$

$$x = (W^{-1} A^T)(A W^{-1} A^T)^{-1} b \quad \longrightarrow \quad \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} = \begin{bmatrix} 7.6190 \\ 1.9048 \\ 0.4762 \end{bmatrix}$$

(3x1) (3x3) (3x1) (1x3)(3x3) (3x1) (1x1)

Following (Handwritten) Slides

- Derivation of least squares solutions
- Singular value decomposition (SVD) – method used by MATLAB pinv
- m-files for examples

$$\min_{\underline{x}} \underline{x}^T \underline{x} \quad \text{s.t. } \underline{\underline{A}}\underline{x} = b$$ Using Lagrange multipliers

$$L(\underline{x}, \underline{\lambda}) = \underline{x}^T \underline{x} + \lambda^T (\underline{\underline{A}}\underline{x} - \underline{b})$$

$$\frac{\partial L}{\partial \underline{x}} = 2\underline{x} + \underline{\lambda}^T \underline{\underline{A}} = \underline{0} \quad \left(\text{also, } \underline{\lambda}^T \underline{\underline{A}} = \underline{\underline{A}}^T \underline{\lambda}\right)$$

$$= 2\underline{x} + \underline{\underline{A}}^T \underline{\lambda} = 0 \implies \underline{x} = -\frac{1}{2} \underline{\underline{A}}^T \lambda$$

$$\frac{\partial L}{\partial \lambda} = \underline{\underline{A}}\underline{x} - \underline{b} = 0 \implies -\frac{1}{2} \underline{\underline{A}}\underline{\underline{A}}^T \underline{\lambda} - \underline{b} = 0$$

$$\underline{\underline{A}}\underline{\underline{A}}^T \lambda = -2b \implies \underline{\lambda} = -2(\underline{\underline{A}}\underline{\underline{A}}^T)^{-1} \underline{b}$$

but $\underline{x} = -\frac{1}{2} \underline{\underline{A}}^T \lambda$, so $\underline{x} = -\frac{1}{2} \underline{\underline{A}}^T (-2)(\underline{\underline{A}}\underline{\underline{A}}^T)^{-1} \underline{b}$

$$\implies \boxed{\underline{x} = \underline{\underline{A}}^T (\underline{\underline{A}} A^T)^{-1} \underline{b}} \stackrel{?}{=} \boxed{(\underline{\underline{A}}^T \underline{\underline{A}})^{-1} \underline{\underline{A}}^T \underline{b}} \quad \checkmark$$ See below

pre & post multiply by A $\implies$ $$\underbrace{\underline{\underline{A}}\underline{\underline{A}}^T (\underline{\underline{A}}\underline{\underline{A}}^T)^{-1}}_{\underline{\underline{I}}} \underline{\underline{A}} \stackrel{?}{=} \underline{\underline{A}} \underbrace{(\underline{\underline{A}}^T \underline{\underline{A}})^{-1} \underline{\underline{A}}^T \underline{\underline{A}}}_{\underline{\underline{I}}} \quad \checkmark$$

$$\min_{x} \; x^T W x \qquad \text{s.t.} \; A x = b$$

$$L(x, \lambda) = x^T W x + \lambda^T (A x - b)$$

$$\frac{\partial L}{\partial x} = 2 W x + A^T \lambda = 0 \implies W x = -\frac{1}{2} A^T \lambda$$

$$x = -\frac{1}{2} W^{-1} A^T \lambda$$

$$\frac{\partial L}{\partial \lambda} = A x - b = 0 \implies -\frac{1}{2} A W^{-1} A^T \lambda - b = 0$$

$$A W^{-1} A^T \lambda = -2 b \implies \lambda = -2 (A W^{-1} A^T)^{-1} b$$

& since $x = -\frac{1}{2} W^{-1} A^T \lambda \implies x = -\frac{1}{2} W^{-1} A^T (-2) (A W^{-1} A^T)^{-1} b$

$$\implies \boxed{x = W^{-1} A^T (A W^{-1} A^T)^{-1} b}$$

SVD Singular Value Decomposition

$$\underline{\underline{A}}_{(n\times m)} = \underline{\underline{U}}_{(n\times n)}\ \underline{\underline{\Sigma}}_{(n\times m)}\ \underline{\underline{V}}^T_{(m\times m)}$$

$\underline{\underline{U}}$ = columns are orthonormal (left singular vector matrix

$$\Rightarrow \underline{\underline{U}}^T_{(n\times n)}\underline{\underline{U}}_{(n\times n)} = \underline{\underline{I}}_{(n\times n)}$$

$\underline{\underline{V}}$ = columns are orthonormal (right sing. vec. matrix

$$\Rightarrow \underline{\underline{V}}^T_{(m\times m)}\underline{\underline{V}}_{(m\times m)} = \underline{\underline{I}}_{(m\times m)}$$

$\underline{\underline{\Sigma}}$ = matrix of singular values

$$\underline{\underline{\Sigma}}_{n\times m} = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & \sigma_n & 0 & 0 & 0 \end{bmatrix}$$

$n < m$

We are solving $\underline{\underline{A}}\,\underline{x} = \underline{b}$, where $\underline{\underline{A}} = \underline{\underline{U}}\,\underline{\underline{\Sigma}}\,\underline{\underline{V}}^T$

$$(\underline{\underline{U}}\,\underline{\underline{\Sigma}}\,\underline{\underline{V}}^T)\underline{x} = \underline{b}$$

$$(\underline{\underline{V}}\,\underline{\underline{\Sigma}}^{-1}\underline{\underline{U}}^T)(\underline{\underline{U}}\,\underline{\underline{\Sigma}}\,\underline{\underline{V}}^T)\,\underline{x} = \underline{\underline{V}}\,\underline{\underline{\Sigma}}^{-1}\underline{\underline{U}}^T\underline{b}$$

$\underline{\underline{U}}^T\underline{\underline{U}} = I$, then $\underline{\underline{\Sigma}}^{-1}\underline{\underline{\Sigma}} = I$, then $\underline{\underline{V}}\,\underline{\underline{V}}^T = I$

$$\Rightarrow \boxed{\underline{x} = \underline{\underline{V}}\,\underline{\underline{\Sigma}}^{-1}\underline{\underline{U}}^T\underline{b}}$$ pinv in MATLAB

for $n \times m$

$$\underline{\underline{\Sigma}}^{-1} = \underbrace{\begin{bmatrix} 1/\sigma_1 & 0 & \cdots & \cdots \\ 0 & 1/\sigma_2 & 0 & \cdots \\ 0 & \cdots & 0 & 1/\sigma_n \\ 0 & \cdots & \cdots & 0 \end{bmatrix}}_{m \times n}$$

% parameter fitting as a least squares problem

```
%
 a = 1;
 b = 1;
%
 t = 0:2:10;   % independent variable (creates row vector)
 t = t';       % creates column vector
%  add some noise
 rng('default')   % same random sequence each time the script is run
 noise = 0.5*randn(6,1);
 z = a + b.*t + noise;   % measured dependent variable
%
 figure(1)             % plot of data
 plot(t,z,'ko')         % black circles for data points
 xlabel('t')           % xaxis label
 ylabel('z')           % yaxis label
 title('Find best fit of z(i) = alpha + beta*t(i)')
%
 A = [ones(6,1) t];
 b = z;
 x = inv(A'*A)*A'*b         % least squares solution for alpha & beta
%
 alpha = x(1);
 beta  = x(2);
%
 zhat = alpha + beta.*t;   % model dependent variable
%
 figure(2)               % compare data and model on same plot
 plot(t,z,'ko',t,zhat,'k')
 legend('data','model')
 xlabel('t')
 ylabel('z')
 title('Best fit of z(i) = alpha + beta*t(i)')
%
```

```
%  alternative to inv in MATLAB
%
 xcheck = (A'*A)\A'*b
%
%  use pinv -- based on SVD
%
 xpinv = pinv(A)*b
%
%  note that all three results yielded the same alpha and beta
```

```
% flow example
% outlet flow, F4, equals sum of inlet flows, F1+F2+F3
%
% minimize the sum F1^2 + F2^2 + F3^2
%  s.t. F1+F2+F3 = F4
% that is, min x'x, s.t. Ax=b
%
  Aflow = [1 1 1]
  bflow = 10
  Aflow'*Aflow
  rank(Aflow'*Aflow)
%  Flowvec = (Aflow'*Aflow)\Aflow'*bflow  (*** not invertible! ***)
%
%  the alternative x = A'(AA')^-1*b is better!
%
  Flowv   = Aflow'*inv(Aflow*Aflow')*bflow
%
%  now, use the SVD based method (pinv)
%
  Flowvec = pinv(Aflow)*bflow   % pseudo-inverse uses SVD
  sum(Flowvec)
%
%  illustrate SVD to calculate the generalized inverse
%
  [U,S,V] = svd(Aflow)
%  using knowledge of dimensions for the next line
  FlowvecSVD = V(:,1)*(1/S(1,1))*U'*bflow
%
```

```
% -------------------------------------
% weighted flow example
% valves sized such that valve 1 can handle 4 times as much
% flow as valve 2
%   which can handle 4 times as much flow as valve 3
% The following weight matrix is then used
  Wflow = [1 0 0; 0 4 0; 0 0 16]
%
% FlowvecW = (Aflow'*Wflow*Aflow)\Aflow'*Wflow*bflow
  Winv = inv(Wflow);
  FlowvecW = Winv*Aflow'*inv(Aflow*Winv*Aflow')*bflow
  sum(FlowvecW)
```