Module – 3.2

STATIC, DETERMINISTIC LINEAR INVERSE PROBLEM ILL-POSED PROBLEM AND IMPERFECT MODEL

S. Lakshmivarahan

School of Computer Science

University of Oklahoma

Norman, Ok – 73069, USA

varahan@ou.edu

WELL AND ILL-POSED PROBLEMS

- Let $Z = Hx$, $Z \in R^m$, $x \in R^n$, $H \in R^{mxn}$

- Module 3.1 contains solution to the well-posed linear least squares problems when the matrix H is full rank, that is Rank(H) = min{m, n}

- If the Rank(H) < min{m, n}, then H is rank-deficient and the problem is ill-posed

- In this case, the Grammian matrices H^TH and HH^T are singular

ILL-POSED PROBLEM – TIKHONOV REGULARIZATION

- When is H is rank deficient, we cannot use the formula

$$H^{+} = (H^{T}H)^{-1}H^{T} \text{ or } H^{+} = H^{T}(HH^{T})^{-1}$$

for the generalized inverse of H in computing X_{LS}

- While we could still compute H^{+} using the method of singular value decomposition (SVD) (Module – 4.2), we seek alternate formulation of the least squares problem
- The method of regularization due to Tikhonov is used to get around the rank deficiency of H

TIKHONOV'S METHOD

- Define

$$f(x) = \alpha\|x\|_2^2 + \|Z - Hx\|_2^2 \qquad \text{-> (1)}$$

- The addition of $\alpha\|x\|_2^2$ to the traditional sum of squared error term helps to avoid the challenges resulting from the rank deficiency of H
- Rewriting

$$f(x) = \alpha x^T x + (Z - Hx)^T(Z - Hx) \qquad \text{-> (2)}$$

It readily follows that (Verify!)

$$\nabla_x f(x) = (H^T H + \alpha I)x - H^T Z$$

$$X_{LS}(\alpha) = (H^T H + \alpha I)^{-1} H^T Z \qquad \text{-> (3)}$$

TIKHONOV'S METHOD

- Since H^TH is singular, by adding a diagonal perturbation αI to H^TH, we can ensure that $(H^TH + \alpha I)$ is non-singular

- One could use the Gershgorin Circle theorem to H^TH to estimate the least value of α that would render $(H^TH + \alpha I)$ non-singular

- If H is of full rank, then we can set $\alpha = 0$ and obtain the known least square solution

A MATRIX IDENTITY

- A well known matrix identity:

$$[A^T B^{-1} A + D^{-1}]^{-1} A^T B^{-1} = DA^T [B + ADA^T]^{-1} \quad \text{-> (4)}$$

- Setting A = H, B = I, $D^{-1} = \alpha I$, this identity becomes:

$$[H^T H + \alpha I]^{-1} H^T = \alpha^{-1} H^T [I + \alpha^{-1} HH^T]^{-1}$$

$$= \alpha^{-1} H^T [\alpha^{-1}(\alpha I + HH^T)]^{-1}$$

$$= H^T [\alpha I + HH^T]^{-1} \quad \text{-> (5)}$$

A UNIFIED APPROACH

- Substituting (5) in (3):

$$X_{LS}(\alpha) = (H^T H + \alpha I)^{-1} H^T Z, \quad m > n \qquad \text{-> (6)}$$

$$= H^T[\alpha I + HH^T]Z, \quad m < n \qquad \text{-> (7)}$$

- Setting $\alpha = 0$ in (6) leads the optimal solution to the full rank problem when $m > n$ – Refer to Module – 3.1
- Setting $\alpha = 0$ in (7) leads to the optimal solution to the full rank problem when $m < n$ – Refer to Module – 3.1

PERFECT VS IMPERFECT MODEL

- The saying goes: “no model is perfect, but some models are useful”
- Often assume that a model is <u>perfect</u>
- Imperfection is a model come from various directions:

  - complete physics, wrong parametrization, etc
- Irrespective of whether the model is perfect or not, in the overdetermined case, the model is <u>inconsistent</u> in the sense we saw in Module 3.1
- In the underdetermined case, the choice of the method depends on whether or not the model is perfect

STRONG VS WEAK CONSTRAINED FORMULATION: m < n

- When m < n, and the model is perfect, we strictly enforce the model constraint using the Lagrangian multiplier method – see Module 3.1 for details
- This is often called the Strong Constraint formulation
- If the model is not perfect, it is pointless to enforce it strictly
- We require the model equation to be satisfied only approximately
- This is known as the Weak Constraint formulation

STRONG CONSTRAINT FORMULATION - REVISITED

- Let $Z = Hx$ with $z \in R^m$, $x \in R^n$, $H \in R^{m x n}$ and $m < n$
- Assume that H is of full rank and recall that there are infinitely many solutions
- Seek an unique solution that minimizes the following cost functional:

$$J(x) = \frac{1}{2} x^T A x - b^T x + c \quad \text{-> (8)}$$

- Strong constraint formulation:

Minimize $L(x, \lambda)$ where $\lambda \in R^m$ and

$$L(x, \lambda) = J(x) + \lambda^T (Z - Hx) \quad \text{-> (9)}$$

STRONG CONSTRAINT - CONTINUED

- The necessary condition are:

$$\nabla_x L(x, \lambda) = 2Ax - b - H^T \lambda = 0 \qquad \text{-> (10)}$$

$$\nabla_\lambda L(x, \lambda) = Z - Hx = 0 \qquad \text{-> (11)}$$

- Express x in terms of λ using (10), substitute in (11), it can be verified that the strong solutions are

$$\left.\begin{aligned} \lambda_s &= (HA^{-1}H^T)^{-1}[Z - HA^{-1}b] \\ X_s &= A^{-1}b + A^{-1}H^T[HA^{-1}H^T]^{-1}[Z - HA^{-1}b] \end{aligned}\right\} \qquad \text{-> (12)}$$

- Setting b = 0, c = 0 and A = I, we get the solution

$$\left.\begin{aligned} \lambda_s &= (HH^T)^{-1}Z \\ X_s &= H^T(HH^T)^{-1}Z \end{aligned}\right\} \qquad \text{-> (11)}$$

as given in Module 3.1

WEAK CONSTRAINED FORMULATION

- Let $\alpha > 0$ and define a Penalty function

$$P_\alpha(x) = J(x) + \frac{\alpha}{2}(Z - Hx)^T(Z - Hx) \qquad \text{-> (14)}$$

- The necessary condition for minimum is given by

$$\nabla_x P_\alpha(x) = Ax - b + \alpha H^T(Hx - Z) = 0 \qquad \text{-> (15)}$$

- Solving:

$$x(\alpha) = x_1(\alpha) + x_2(\alpha) \qquad \text{-> (16)}$$

$$x_1(\alpha) = (A + \alpha H^T H)^{-1} b \qquad \text{-> (17)}$$

$$x_2(\alpha) = \alpha(A + \alpha H^T H)^{-1} H^T Z \qquad \text{-> (18)}$$

SHERMAN – MORRISON – WOODBURY (SMW) FORMULA

- Let $H \in R^{mxn}$, $\varepsilon_x \in R^{nxn}$, $\varepsilon_V \in R^{mxm}$ $m < n$
- SMW Formula – (two versions)
- $[H^T\varepsilon_v^{-1}H + \varepsilon_v^{-1}]^{-1} = \varepsilon_x - \varepsilon_x H^T[H\varepsilon_x H^T + \varepsilon_v]^{-1}H\varepsilon_x$ -> (19)
- $[H\varepsilon_x H^T + \varepsilon_v]^{-1} = \varepsilon_v^{-1} - \varepsilon_v^{-1}H[H^T\varepsilon_v^{-1}H + \varepsilon_v^{-1}]^{-1}H^T\varepsilon_v$ -> (20)
- Multiplying both side of (20) on the left by $\varepsilon_x H^T$ and simplifying (refer to LLD (2006) – Chapter 17), obtain the matrix identity

$$\varepsilon_x H^T[H\varepsilon_x H^T + \varepsilon_v]^{-1} = [H^T\varepsilon_v^{-1}H + \varepsilon_v^{-1}]H^T\varepsilon_v^{-1} \quad \text{-> (21)}$$

RELATING WEAK AND STRONG SOLUTION

- Applying (20) to (17) with ε_x^{-1} = A, ε_v^{-1}= αIm:

$$(A + \alpha HH^T)^{-1} = A^{-1} - A^{-1}H^T[HA^{-1}H + \alpha^{-1}\mathrm{Im}]^{-1}HA^{-1}$$

$$\to A^{-1} - A^{-1}H^T[HA^{-1}H]^{-1}HA^{-1} \text{ as } \alpha \to \infty \qquad \to (22)$$

- Hence, from (17)

$$X_1^* = \lim_{\alpha \to \infty} x_1(\alpha) = A^{-1}b - A^{-1}H^T(HA^{-1}H^T)^{-1}HA^{-1}b \qquad \to (23)$$

RELATING WEAK AND STRONG SOLUTIONS

- Applying (21) to (18) with ε_x^{-1} = A, ε_v^{-1}= αIm:

$$\alpha(A + \alpha HH^T)^{-1}H^T = A^{-1}H^T[HA^{-1}H + \alpha^{-1}\mathrm{Im}]^{-1}$$

$$\to A^{-1}H^T[HA^{-1}H]^{-1} \text{ as } \alpha \to \infty \qquad \to (24)$$

- Hence, from (18)

$$X_2^* = \lim_{\alpha \to \infty} x_2(\alpha) = A^{-1}H^T(HA^{-1}H^T)^{-1}Z \qquad \to (25)$$

- $$X_1^* + X_2^* = A^{-1}b + A^{-1}H^T(HA^{-1}H^T)^{-1}[Z - HA^{-1}b] \qquad \to (26)$$

$$= X_s \text{ in } (12)$$

- That is, in the limit as the penalty parameter α increases without bund, the weak solution converges to the strong solution

EXERCISES

(7.1) Compute the Gradient and Hessian

$$f(x) = \alpha x^T x + (Z - Hx)^T (Z - Hx)$$

and verify the relation (3)

(7.2) Verify that (12) gives the solution of (10) and (11)

(7.3) Let $H = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1+\varepsilon \end{bmatrix}$ with $\varepsilon > 0$

Compute the eigenvalues of $H^T H$ and plot them as a function of ε for

$-1 \le \varepsilon \le 1$

EXERCISES

(7.4) Let A = $\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$, b = $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, H = $\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$, x = $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and z = $(3, 4, 5)^T$

a) Compute the unique minizer x* of $J(x) = \frac{1}{2}x^TAx - b^Tx$ when Z = Hx using Lagrangian multiplier

b) Compute the unique minimizer x(α) of

$$P_\alpha(x) = J(x) + \frac{\alpha}{2}(Z - Hx)^T(Z - Hx) \text{ as a function of } \alpha$$

c) Plot the norm of x(α) Vs α and show $\lim_{\alpha \to \infty} x(\alpha) = x^*$

REFERENCES

- This module follows from Chapter 5 LLD (2006) and the following report

S.Lakshmivarahan (2015) “On the convergence of class of weak solution to the strong solution of an equality constraint minimization. Problem: A direct proof using matrix identities, School of Computer Science, University of Oklahoma, Norman, Ok - 73019 USA