Module – 4.1

INTERLUDE
AN ASSESSMENT AND A WAY FORWARD

S. Lakshmivarahan

School of Computer Science

University of Oklahoma

Norman, Ok – 73069, USA

varahan@ou.edu

LEAST SQUARES FORMULATIONS – A CLASSIFICATION

- Linear vs nonlinear
- Ordinary vs weighted
- Orthogonal vs oblique projection
- Full rank vs rank deficient
- Over vs under determined
- Off-line vs on-line formulations

PATHWAYS TO THE SOLUTION

Two ways

- Directly minimize f(x)
- Iterative minimization
- Module – 4.3

- Solve ∇_xf(x) = 0
- Check if ∇_x^2f(x) is SPD
- Matrix methods
- Module – 4.2

- These methods are useful throughout the course

PATHWAYS TO SOLUTION – LINEAR CASE

- Off-line, ordinary linear, full rank formulation: Module – 3.1

  $(H^T H)x = H^T Z$, $(H^T H)$ – SPD, $m > n$

  $(HH^T)y = Z$ and $x = H^T y$, (HH^T) – SPD, $m < n$

- Off-line, weighted, linear full rank formulation: Module – 3.1

  $(H^T WH)x = H^T WZ$, $(H^T WH)$ – SPD, $m > n$

  $(HWH^T)y = Z$ and $x = WH^T y$, (HWH^T) – SPD, $m < n$

- W - SPD

PATH WAYS TO SOLUTION – LINEAR CASE

- Off-line, ordinary linear, rank-deficient formulation: Module – 3.2

  $(H^TH + \alpha I)x = H^TZ$, $(H^TH + \alpha I)$ – SPD, $\alpha > 0$, $m > n$

  $(HH^T + \alpha I)y = Z$ and $x = H^Ty$, $(HH^T + \alpha I)$ – SPD, $\alpha > 0$, $m < n$

- On-line, ordinary linear, full rank case: Module – 10

  $X^*(m+1) = X^*(m) + K_{m+1}h_{m+1}[z_{m+1} - h_{m+1}^T x^*(m)]$

  $K_m^{-1} = H^TH$

  $K_{m+1}^{-1} = K_m^{-1} + h_{m+1}h_{m+1}^T$

- K_{m+1} is obtained by Sherman-Morrison-Woodbury formula

PATH WAYS TO SOLUTION – NONLINEAR CASE

- Off-line, ordinary, nonlinear case: Module – 3.4
  - Solve a set of non-linear equations:

$$\nabla_x f(x) = D_x^T(h)[h(x) - Z] = 0$$

- On-line, ordinary, nonlinear case:
  - Use first-order approximation to h(x)

$$[D_{x_c}^T(h)D_{x_c}(h)](x - x_c) = D_{x_c}^T(h)Z$$

$$x_c^{new} \leftarrow x_c + (x - x_c)$$

  - Use second-order approximation to h(x)

$$[D_{x_c}^T(h)D_{x_c}(h) - \sum_{i=1}^{m} g_i \nabla_x^2 h_i(x_c)](x - x_c) = D_{x_c}^T(h)Z$$

$$g = z - h(x_c),\ x_c^{new} \leftarrow x_c + (x - x_c)$$

A WAY FORWARD – COMING ATTRACTIONS:

- Method of normal equations
- At its core, need methods for solving

  $(H^T H)x = H^T Z$ or $(HH^T)y = Z$
- These are linear systems of the form

  $Ax = b$

  with A as a symmetric positive definite (SPD) matrix
- A stand method to linear systems with SPD matrix is the <u>Cholesky decomposition</u> method – (Module 4.2)

ALTERNATE DECOMPOSITION (QR AND SVD) METHODS

- Instead of decomposing (H^TH) or (HH^T) using cholesky method, we can directly decompose H that will simplify the form of the least squares solution

- QR – decomposition – (Module 4.2)

- (SVD) – (Module 4.2)

DIRECT MINIMIZATION OF f(x)

- Use well known iterative methods to directly minimize f(x): (Module 4.3):
  - Gradient method
  - Conjugate gradient method
  - Quasi-Newton methods

REFERENCES

- This Module is a summary of Chapters 5 -7 in LLD (2006)