12.1

Linear equality constrained Min in R^n

Katta G. Murty, IOE 611 Lecture slides

Consider: $\min \theta(x)$ s. to $Ax = b$

where $A_{m\times n}$ has rank m.

Let $\bar{x}$ be any feasible solution. Then the general solution is $\bar{x} + \eta$, where η is a homogeneous sol., i.e., $A\eta = 0$.

Let B be any basis, and $(B{:}D)$ the basic, nonbasic partition for A. Let

$$Z_{n\times(n-m)} = \begin{pmatrix} -B^{-1}D \\ I_{n-m} \end{pmatrix}$$

Then the general homogeneous sol. is of form $Z\xi$ where $\xi = (\xi_1, \ldots, \xi_{n-m})^T \in R^{n-m}$. So, the general feasible sol. is $\bar{x} + Z\xi$.

Our problem equivalent to unconstrained min of $f(\xi) = \theta(\bar{x} + Z\xi)$ over $\xi \in R^{n-m}$.

$\nabla_\xi f(\xi) = \nabla_x \theta(\bar{x} + Z\xi)Z$, called **reduced gradient**.

$\nabla^2_{\xi\xi} f(\xi) = Z^T H(\bar{x} + Z\xi)Z$, called **reduced Hessian**, where

$H(x) = \nabla^2_{xx}\theta(x)$.

1st order nec. opt. conds.: $\nabla\theta(x)Z = 0$.

2nd. order nec. opt. conds.: $Z^T H(x)Z$ is PSD.

Descent Methods: Generate descent sequence $x^0, x^1, \dots$ All descent directions satisfy $A\eta = 0$, so are feasible directions for any step length.

When current pt. is x^r,

Step 1: Search direction in x-space is $\eta^r = Z\xi^r$ where $\xi^r \in R^{n-m}$ is search direction for $f(\xi) = \theta(x^r + Z\xi)$ at $\xi = 0$ in ξ-space.

Steepest descent: Uses $\xi^r = -(\nabla\theta(x^r)Z)^T$.

Newton: Uses ξ^r which is a sol. of $(Z^T H(x^r)Z)\xi = -Z^T(\nabla\theta(x^r))^T$.

Modified Newton: Replace coeff. matrix of ξ in LHS of above by $(Z^T H(x^r)Z) + \gamma I_{n-m}$ where $\gamma > 0$ is suitably choosen.

BFGS: Uses $\xi^r = -(B_r)^{-1}Z^T(\nabla\theta(x^r))^T$ where B_r of order $(n-m)\times(n-m)$ is obtained by using $B_0 = I_{n-m}$, and by updating using the updating formula given earlier, with $s^r = Z^T(x^{r+1}-x^r)$, $y^r = Z^T(\nabla\theta(x^{r+1})-\nabla\theta(x^r))^T$.

C. Grad. : Search direction in x-space at x^0 is $\eta^0 = -Z(\nabla\theta(x^0)Z)^T$ and at x^r is $\eta^r = -Z(\nabla\theta(x^r)Z)^T + \beta_r\eta^{r-1}$, where β_r here are obtained by replacing all $\nabla\theta(x)$ in fomulae given earlier by $\nabla\theta(x)Z$.

Step 2: Step length in x-space is determined by any of the methods discussed earlier.

Finite difference approx. of Hessian: Reduced Hessian $Z^T H(x)Z$ can be approximated. Let ϵ_i be finite difference interval. For $i = 1$ to $n-m$, let $W_{.i} = \frac{1}{\epsilon_i}(\nabla\theta(x+\epsilon_i Z_{.i}) - \nabla\theta(x))^T$ and let $W_{n\times(n-m)}$ be the matrix with $W_{.i}$ as columns. Then an approx. of $Z^T H(x)Z$ is $\frac{1}{2}(Z^T W + W^T Z)$.

Legrange Multiplier Vector: Given an approx. to a local min $\bar{x}$, the associated Lagrange multiplier vector π can be determined to min $||(\nabla\theta(\bar{x}))^T - \pi A||$. It is:

$$\bar{\pi} = (AA^T)^{-1}A(\nabla\theta(\bar{x}))^T$$