10.1

Nonlinear Equations

Katta G. Murty, IOE 611 Lecture slides

If no. of eqs. $>$ [$<$] no. of variables system called **overdetermined** [**underdetermined**] system.

We consider **Square system** of n eqs. in n unknowns, $F(x) = (f_1(x), \dots, f_n(x)) = 0$.

Newton's Method (or Newton-Raphson Method)

This is the iteration:

$$x^{r+1} = x^r - (\nabla_x F(x^r))^{-1} F(x^r)$$

assuming that $\nabla_x F(x^r)$ is nonsingular.

Examples: 1) $x_1^2 + x_2^2 - 1 = 0$, $x_1^2 - x_2 = 0$, $\quad x^0 = (1, 0)^T$.

2) $x_1 + x_2 - 3 = 0$, $x_1^2 + x_2^2 - 9 = 0$, $\quad x^0 = (1, 4)^T$.

Theorem: Local convergence of Newton's Method: Suppose there exists x^* s. th. $F(x^*) = 0$, and $\nabla_x F(x^*)$ is nonsingular and $||(\nabla_x F(x^*))^{-1}|| \leq \beta$ for some $\beta > 0$. Also suppose that $\nabla_x F(x)$ is Lipschitz continuous with constant γ. Then there exists an open nbhd. of x^* s. th. $\forall x^0$ in this nbhd. the sequence $\{x^r\}$ generated by Newton's method converges to x^* and obeys for $r = 0, 1, \ldots$

$$||x^{r+1} - x^*|| \leq \beta\gamma||x^r - x^*||^2$$

Broyden's Method

Most popular secant method for solving nonlinear eqs. It approximtes $\nabla_x F(x)$.

Initiated with some x^0 and $B_0 = \nabla_x F(x^0)$. General iteration is:

$$x^{r+1} = x^r - B_r^{-1} F(x^r), \quad \text{where}$$

$$\text{Updating formula} \quad B_{r+1} = B_r + \frac{(y^r - B_r s^r)(s^r)^T}{(s^r)^T s^r}, \; r = 0, 1, \ldots$$

$$y^r = F(x^{r+1}) - F(x^r), \quad s^r = x^{r+1} - x^r$$

It can be shown to locally converge superlinearly under same conds. as Newton's method. Requires less function evaluations than finite difference Newton.

When implemented, instead of using the updating formula for B_r, an equivalent updating formula for QR-factorization of B_r is used so that s^{r+1} can be computed using only $O(n^2)$ effort.

Both methods are locally convergent. Globally convergent methods for $F(x) = 0$ are derived thro' unconstrained min of $(F(x))^T F(x)$.

Affine scaling method for nonlinear eqs. with bounds on vars.

Consider solving:

$$f_i(x) = b_i, \quad i = 1 \text{ to } m.$$

$$x \in \Gamma = \{x : \quad x_j \geq \alpha_j \text{ for } j \in J_1; \quad x_j \leq \beta_j \text{ for } j \in J_2;$$

$$\alpha_j \leq x_j \leq \beta_j \text{ for } j \in J_3; \quad x_j \text{ unrestricted for } j \in J_4 \}$$

where α_j, β_j are reals and for $j \in J_3$, $\alpha_j < \beta_j$; and (J_1, J_2, J_3, J_4) is a partition of $\{1, \ldots, n\}$.

Starting point $x^0 \in$ Interior(Γ), and all iterates will be in interior(Γ). Let $F(x) = (f_i(x) : i = 1 \text{ to } m)^T$.

Given $x^k \in$ interior(Γ), define weight vector $\sigma^k = (\sigma_1^k, \ldots, \sigma_n^k)^T$ for it by

$$\sigma_j^k = \begin{cases} (x_j^k - \alpha_j)^2 & j \in J_1 \\ (\beta - x_j^k)^2 & j \in J_2 \\ \min\{(x_j^k - \alpha_j)^2, (\beta_j - x_j^k)^2\} & j \in J_3 \\ N_j > 0 & j \in J_4 \end{cases}$$

Define

$$D_k = \text{diag}(\sigma_k)$$

$$r^k = \text{residual vector } (b_i - f_i(x^k) : i = 1 \text{ to } m)^T.$$

$$B^k_{m \times m} = \nabla_x F(x^k) D_k (\nabla_x F(x^k))^T$$

Let $u^k = (u^k_1, \ldots, u^k_m)^T$ be the sol. to $B^k u = r^k$. If Jacobian has full row rank and $\sigma^k > 0$, this system has unique sol. which is the minimizer of $\frac{1}{2} \Sigma_{j=1}^n \sigma^k_j (((\nabla_x F(x^k))_{.j})^T u)^2 - (r^k)^T u$.

The correction direction at x^k is $s^k \in R^n$ determined to minimize $\frac{1}{2} \Sigma_{j=1}^n s_j^2 / \sigma_j^k$ s. to $(\nabla_x F(x^k)) s = r^k$. For this problem, u^k is the opt. Lagrange multiplier vector, and s^k itself is given by

$$s^k = D_k \delta^k \quad \text{where} \quad \delta^k = (\nabla_x F(x^k))^T u^k$$

Choose the new pt. to be $x_{k+1} = x^k + \lambda_k s^k$ where $\lambda_k = \rho \mu_k$; $0 < \rho < 1$ and μ_k is the maximum step length λ that keeps $x^k + \lambda s^k$ within Γ.

Terminate when either $||r^k||$ is small; or when step length λ_k becomes close to 0.

Reference: I. I. Dikin, "Determination of Interior Points of Systems of Inequality and Equality Constraints", 1997.