Notes for Part 5: Penalty and augmented Lagrangian methods for equality constrained optimization

Nick Gould, CSED, RAL, Chilton, OX11 0QX, England (n.gould@rl.ac.uk)

January 31, 2006

5 Sketches of proofs for Part 5

5.1 Proof of Theorem 5.1

Denote the left generalized inverse of $A^T(x)$ by

$$A^+(x) = \Big(A(x)A^T(x)\Big)^{-1} A(x)$$

at any point for which $A(x)$ is full rank. Since, by assumption, $A(x_*)$ is full rank, these generalized inverses exists, and are bounded and continuous in some open neighbourhood of x_*.

Now let

$$y_k = -\frac{c(x_k)}{\mu_k}$$

as well as

$$y_* = A^+(x_*)g(x_*).$$

It then follows from the inner-iteration termination test

$$\|g(x_k) - A^T(x_k)y_k\| \le \epsilon_k \tag{5.1}$$

and the continuity of $A^+(x_k)$ that

$$\|A^+(x_k)g(x_k) - y_k\|_2 = \Big\|A^+(x_k)\Big(g(x_k) - A^T(x_k)y_k\Big)\Big\|_2 \le 2\|A^+(x_*)\|_2\epsilon_k.$$

Then

$$\|y_k - y_*\|_2 \le \|A^+(x_*)g(x_*) - A^+(x_k)g(x_k)\|_2 + \|A^+(x_k)g(x_k) - y_k\|_2$$

which implies that $\{y_k\}$ converges to y_*. In addition, continuity of the gradients and (5.1) implies that

$$g(x_*) - A^T(x_*)y_* = 0,$$

while the fact that $c(x_k) = -\mu_k y_k$ with bounded y_k implies that

$$c(x_*) = 0.$$

Hence (x_*, y_*) satisfies the first-order optimality conditions.

5.2 Proof of Theorem 5.2

The proof of convergence of y_k to $y_* \stackrel{\text{def}}{=} A^+(x_*)g(x_*)$ for which $g(x_*) = A^T(x_*)y_*$ is exactly as for Theorem 5.1. For the second part of the theorem, the definition of y_k and the triangle inequality gives

$$\|c(x_k)\| = \mu_k\|u_k - y_k\| \le \mu_k\|y_k - y_*\| + \mu_k\|u_k - y_*\|.$$

the first term on the right-hand side converges to zero as y_k approaches y_* with bounded μ_k, while the second term has the same limit because of the assumptions made. Hence $c(x_*) = 0$, and (x_*, y_*) satisfies the first-order optimality conditions.