

## Notes for Part 5: SQP methods for equality constrained optimization

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

January 4, 2006

### 5 Sketches of proofs for Part 5

#### 5.1 Proof of Theorem 5.1

The SQP search direction  $s_k$  and its associated Lagrange multiplier estimates  $y_{k+1}$  satisfy

$$B_k s_k - A_k^T y_{k+1} = -g_k \quad (5.1)$$

and

$$A_k s_k = -c_k. \quad (5.2)$$

Premultiplying (5.1) by  $s_k$  and using (5.2) gives that

$$s_k^T g_k = -s_k^T B_k s_k + s_k^T A_k^T y_{k+1} = -s_k^T B_k s_k - c_k^T y_{k+1} \quad (5.3)$$

Likewise (5.2) gives

$$\frac{1}{\mu_k} s_k^T A_k^T c_k = -\frac{\|c_k\|^2}{\mu_k}. \quad (5.4)$$

Combining (5.3) and (5.4), and using the positive definiteness of  $B_k$ , the Cauchy-Schwarz inequality and the fact that  $s_k \neq 0$  if  $x_k$  is not critical, yields

$$s_k^T \nabla_x \Phi(x_k) = s_k^T \left( g_k + \frac{1}{\mu_k} A_k^T c_k \right) = -s_k^T B_k s_k - c_k^T y_{k+1} - \frac{\|c_k\|^2}{\mu_k} < -\|c_k\|_2 \left( \frac{\|c_k\|_2}{\mu_k} - \|y_{k+1}\|_2 \right) \leq 0$$

because of the required bound on  $\mu_k$ .

#### 5.2 Proof of Theorem 5.2

The proof is slightly complicated as it uses the calculus of non-differentiable functions. See Theorem 14.3.1 in

R. Fletcher, "Practical Methods of Optimization", Wiley (1987, 2nd edition),

where the converse result that if  $x_*$  is an isolated local minimizer of  $\Phi(x, \rho)$  for which  $c(x_*) = 0$ , then  $x_*$  solves the given nonlinear program so long as  $\rho$  is sufficiently large, is also given. Moreover, Fletcher shows (Theorem 14.3.2) that  $x_*$  cannot be a local minimizer of  $\Phi(x, \rho)$  when  $\rho < \|y_*\|_D$ .

### 5.3 Proof of Theorem 5.3

For small steps  $\alpha$ , Taylor's theorem applied separately to  $f$  and  $c$ , along with (5.2), gives that

$$\begin{aligned} \Phi(x_k + \alpha s_k, \rho_k) - \Phi(x_k, \rho_k) &= \alpha s_k^T g_k + \rho_k \left( \|c_k + \alpha A_k s_k\| - \|c_k\| \right) + O(\alpha^2) \\ &= \alpha s_k^T g_k + \rho_k \left( \|(1 - \alpha)c_k\| - \|c_k\| \right) + O(\alpha^2) \\ &= \alpha \left( s_k^T g_k - \rho_k \|c_k\| \right) + O(\alpha^2) \end{aligned}$$

Combining this with (5.3), and once again using the positive definiteness of  $B_k$ , the Hölder inequality and the fact that  $s_k \neq 0$  if  $x_k$  is not critical, yields

$$\begin{aligned} \Phi(x_k + \alpha s_k, \rho_k) - \Phi(x_k, \rho_k) &= -\alpha \left( s_k^T B_k s_k + c_k^T y_{k+1} + \rho_k \|c_k\| \right) + O(\alpha^2) \\ &< -\alpha \left( -\|c_k\| \|y_{k+1}\|_D + \rho_k \|c_k\| \right) + O(\alpha^2) \\ &= -\alpha \|c_k\| \left( \rho_k - \|y_{k+1}\|_D \right) + O(\alpha^2) < 0 \end{aligned}$$

because of the required bound on  $\rho_k$ , for sufficiently small  $\alpha$ . Hence sufficiently small steps along  $s_k$  from non-critical  $x_k$  reduce  $\Phi(x, \rho_k)$ .