

Here x is the starting point for Newton's method, and $p^* = \inf_x f(x)$ is the optimal value. The constant γ depends only on the backtracking parameters α and β , and is given by

$$\frac{1}{\gamma} = \frac{20 - 8\alpha}{\alpha\beta(1 - 2\alpha)^2}.$$

The constant c depends only on the tolerance ϵ_{nt} ,

$$c = \log_2 \log_2(1/\epsilon_{\text{nt}}),$$

and can reasonably be approximated as $c = 6$. The expression (11.24) is a quite conservative bound on the number of Newton steps required, but our interest in this section is only to establish a complexity bound, concentrating on how it increases with problem size and algorithm parameters.

In this section we use this result to derive a bound on the number of Newton steps required for one outer iteration of the barrier method, *i.e.*, for computing $x^*(\mu t)$, starting from $x^*(t)$. To lighten the notation we use x to denote $x^*(t)$, the current iterate, and we use x^+ to denote $x^*(\mu t)$, the next iterate. We use λ and ν to denote $\lambda^*(t)$ and $\nu^*(t)$, respectively.

The self-concordance assumption implies that

$$\frac{\mu t f_0(x) + \phi(x) - \mu t f_0(x^+) - \phi(x^+)}{\gamma} + c \quad (11.25)$$

is an upper bound on the number of Newton steps required to compute $x^+ = x^*(\mu t)$, starting at $x = x^*(t)$. Unfortunately we do not know x^+ , and hence the upper bound (11.25), until we actually compute x^+ , *i.e.*, carry out the Newton algorithm (whereupon we know the *exact* number of Newton steps required to compute $x^*(\mu t)$, which defeats the purpose). We can, however, derive an upper bound on (11.25), as follows:

$$\begin{aligned} & \mu t f_0(x) + \phi(x) - \mu t f_0(x^+) - \phi(x^+) \\ &= \mu t f_0(x) - \mu t f_0(x^+) + \sum_{i=1}^m \log(-\mu t \lambda_i f_i(x^+)) - m \log \mu \\ &\leq \mu t f_0(x) - \mu t f_0(x^+) - \mu t \sum_{i=1}^m \lambda_i f_i(x^+) - m - m \log \mu \\ &= \mu t f_0(x) - \mu t \left(f_0(x^+) + \sum_{i=1}^m \lambda_i f_i(x^+) + \nu^T (Ax^+ - b) \right) - m - m \log \mu \\ &\leq \mu t f_0(x) - \mu t g(\lambda, \nu) - m - m \log \mu \\ &= m(\mu - 1 - \log \mu). \end{aligned}$$

This chain of equalities and inequalities needs some explanation. To obtain the second line from the first, we use $\lambda_i = -1/(t f_i(x))$. In the first inequality we use the fact that $\log a \leq a - 1$ for $a > 0$. To obtain the fourth line from the third, we use $Ax^+ = b$, so the extra term $\nu^T (Ax^+ - b)$ is zero. The second inequality follows