CNAc: Continuous Optimization
Solutions to problem set 3 — trust-region methods

Honour School of Mathematics, Oxford University
Hilary Term 2006, Dr Nick Gould

Problem 1.

The Cauchy point is $s_k^{\mathrm{C}} = -\alpha_k^{\mathrm{C}} g_k$ where

$$\alpha_k^{\mathrm{C}} = \underset{0<\alpha\leq\Delta_k/\|g_k\|}{\arg\min}\; m_k(-\alpha g_k).$$

To find α_k^{C}:

- if $g_k^T B_k g_k \leq 0$,
$$\alpha_k^{\mathrm{C}} = \frac{\Delta_k}{\|g_k\|}.$$
- otherwise,
$$\alpha_k^{\mathrm{C}} = \min\left(\frac{\|g_k\|_2^2}{g_k^T B_k g_k}, \frac{\Delta_k}{\|g_k\|}\right).$$

Problem 2.

(a) The unconstrained minimizer $-(1,0,1/2)^T$ has ℓ_2-norm $1 < \sqrt{5}/2 < 2$. Thus, since B is positive definite, the unconstrained minimizer solves the problem.

(b) The unconstrained minimizer has too large a ℓ_2-norm, so the solution must lie on the boundary of the constraint. The solution must be of the form $-(1/(1+\lambda),0,1/(2+\lambda))^T$. To satisfy the trust-region constraint, we then must have
$$\frac{1}{(1+\lambda)^2} + \frac{1}{(2+\lambda)^2} = \Delta^2 = \frac{25}{144}$$
which has a root $\lambda = 2$. Thus the required solution is $-(1/3,0,1/4)^T$.

(c) 1(c). The Hessian is indefinite so the solution must lie on the boundary of the constraint. The solution is then of the form $-(1/(-2+\lambda),0,1/(-1+\lambda))^T$. To satisfy the trust-region constraint, we then must have
$$\frac{1}{(-2+\lambda)^2} + \frac{1}{(-1+\lambda)^2} = \Delta^2 = \frac{25}{144}$$
which has a root $\lambda = 5$ (c.f. the previous equation with a change of variables $\hat{\lambda} = \lambda + 3$) at which $B + \lambda I$ is positive semi-definite. Thus again the solution is $-(1/3,0,1/4)^T$.

(d) Again B is indefinite, and so the solution must be of the form $-(\omega,0,1/(-1+\lambda))^T$, where $\omega = 0/(-2+\lambda)$ can only be nonzero if $\lambda = 2$—note that $B+\lambda I$ is only positive semi-definite when $\lambda \geq 2$. Suppose that $\lambda > 2$. To satisfy the trust-region constraint, we then must have
$$\frac{1}{(-1+\lambda)^2} = \Delta^2 = \frac{1}{4}$$
which has roots 1 ± 2. The desired root is $\lambda = 3$, from which we deduce the solution is $-(0,0,1/2)^T$.

(e) As in (d), if we guess that $\lambda > 2$, we find that the roots of

$$\frac{1}{(-1+\lambda)^2} = \Delta^2 = 2$$

are $1 \pm 1/\sqrt{2} < 2$. So λ must be 2, and the solution is of the form $-(\omega, 0, 1)^T$. To satisfy the trust-region constraint, we then must have

$$\omega^2 + 1 = \Delta^2 = 2,$$

and hence $\omega = \pm 1$. Thus the required solution is $-(\pm 1, 0, 1)^T$.

Problem 3.

The problem is convex, so the required solution will satisfy $(B + \lambda I)s = -g$ for some $\lambda \geq 0$. Thus

$$s = -\begin{pmatrix} \frac{1}{1+\lambda} \\ \frac{1}{3+\lambda} \end{pmatrix}.$$

We illustrate this trajectory in Figure 1.

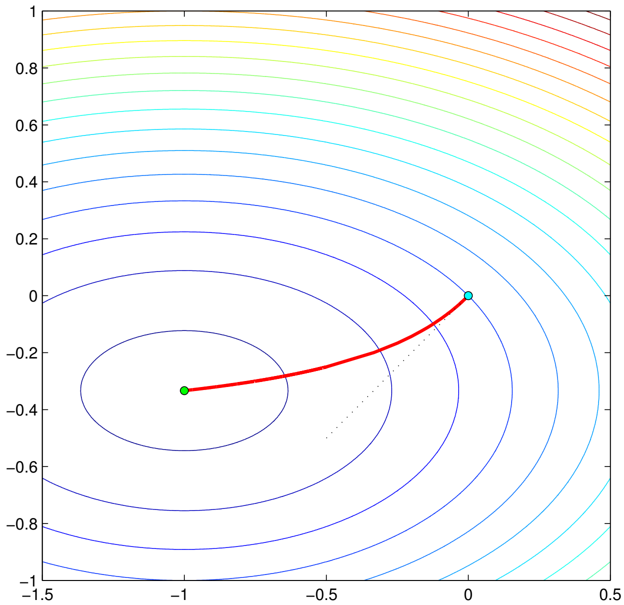

Figure 1: The trust-region minimizer of the quadratic function $0.5s_1^2 + 1.5s_2^2 + s_1 + s_2$ as the trust-region radius increases from zero to infinity. Note that the tangent to the trajectory when the radius is zero is in the steepest-descent direction.

The Lagrange multiplier and the trust-region radius satisfy the relationship

$$\frac{1}{(1+\lambda)^2} + \frac{1}{(3+\lambda)^2} = \Delta^2 \tag{1}$$

for $\lambda \geq 0$. Letting $\theta = \lambda + 2$, (1) implies

$$\Delta^2(\theta^2 - 1)^2 = 2\theta^2 + 2.$$

from which trivial calculations show that

$$\theta = \frac{\sqrt{1 + \Delta^2 + \sqrt{1 + 4\Delta^2}}}{\Delta},$$

i.e.,

$$\lambda = -2 + \frac{\sqrt{1 + \Delta^2 + \sqrt{1 + 4\Delta^2}}}{\Delta}.$$

Using (1), the solution becomes unconstrained when $\lambda = 0$, that is when $\Delta = \sqrt{10}/3$.

Problem 4.

Let $\bar{s} = Rs$. Then the problem we wish to solve is equivalent to

$$\underset{\bar{s} \in \mathbb{R}^n}{\text{minimize}} \quad \bar{s}^T \bar{g} + \tfrac{1}{2} \bar{s}^T \bar{B} \bar{s} \;\; \text{subject to} \;\; \|\bar{s}\|_2 \leq \Delta. \tag{2}$$

where $\bar{g} = R^{-T} g$ and $\bar{B} = R^{-T} H R^{-1}$. But then Theorem 3.9 gives that the global minimizer $\bar{s}_*$ of (2) satisfies

$$(\bar{B} + \lambda_* I)\bar{s}_* = -\bar{g}, \tag{3}$$

where $\bar{B} + \lambda_* I$ is positive semi-definite, $\lambda_* \geq 0$ and $\lambda_*(\|\bar{s}_*\|_2 - \Delta) = 0$. Equation (3) is equivalent to

$$(B + \lambda_* M)s_* = -g, \tag{4}$$

while

$$\bar{B} + \lambda_* I = R^{-T}(B + \lambda_* M)R^{-1}$$

which is positive semi-definite if and only if $B + \lambda_* M$ is. Thus the required conditions are that s_* satisfies (4) where $B + \lambda_* M$ is positive semi-definite, $\lambda_* \geq 0$ and $\lambda_*(\|s_*\|_M - \Delta) = 0$.