$$\lim_{k\to\infty} \left\| H^k - \nabla^2 f(x^k) \right\| = 0, \qquad \lim_{k\to\infty} \frac{\left\| H^k d^k + \nabla f(x^k) \right\|}{\left\| \nabla f(x^k) \right\|} = 0.$$

Show that $\{\|x^k - x^*\|\}$ converges superlinearly.

1.4.6 (www)

Apply Newton's method with a constant stepsize to minimization of the function $f(x) = \|x\|^3$. Identify the range of stepsizes for which convergence is obtained, and show that it includes the unit stepsize. Show that for any stepsize within this range, the method converges linearly to $x^* = 0$. Explain this fact in light of Prop. 1.4.1.

1.4.7

Consider Newton's method with the trust region implementation for the case of a positive definite quadratic cost function. Show that the method terminates in a finite number of iterations.

1.4.8

Generated by CamScanner

Consider Newton's method with the trust region implementation for the case of a positive definite quadratic cost function. Show that the method terminates in a finite number of iterations.

O **1.4.8**

(a) Consider the ~~pure form of Newton's method~~ for the case of the cost function $f(x) = \|x\|^\beta$, where $\beta > 1$. For what starting points and values of β does the method converge to the optimal solution? What happens when $\beta \leq 1$?

(b) Repeat part (a) for the case where Newton's method with the Armijo rule is used.

1.4.9 (Computational Problem)

Consider a firm wishing to maximize its earnings by optimally choosing the selling price of its product, denoted y, and the amount spent for advertising, denoted z. Assume the following relationships:

$$E = yx - \big(z + g_2(x)\big),$$

$$x = g_1(y, z) = a_1 + a_2 y + a_3 z + a_4 yz + a_5 z^2,$$

Generated by CamScanner

Hint: Use the line of argument of Prop. 1.3.2 together with the capture theorem (Prop. 1.2.5). Alternatively, instead of using the capture theorem, consult the proof of the subsequent Prop. 1.4.1.

1.3.3

Consider a positive definite quadratic problem with Hessian matrix Q. Suppose we use scaling with the diagonal matrix whose ith diagonal element is q_{ii}^{-1}, where q_{ii} is the ith diagonal element of Q. Show that if Q is 2×2, this diagonal scaling improves the condition number of the problem and the convergence rate of steepest descent. (*Note*: This need not be true for dimensions higher than 2.)

Proof: let $u^k = x^k - x^*$, thus $u^{k+1} = x^{k+1}$

Generated by CamScanner

q_{ii} is the ith diagonal element of Q. Show that if Q is 2×2, this diagonal scaling improves the condition number of the problem and the convergence rate of steepest descent. (*Note*: This need not be true for dimensions higher than 2.)

Proof: let $u^k = x^k - x^*$, thus $u^{k+1} = x^{k+1} - x^*$,

1.3.4 (Steepest Descent with Errors)

$f(u) = \frac{1}{2} u'Qu$

$x^{k+1} = x^k - s(\nabla f(x^k) + e^k)$

Consider the steepest descent method

$\Rightarrow u^{k+1} = u^k - s(Qu^k + e^k) = (I - sQ)u^k - se^k$

$$x^{k+1} = x^k - s\big(\nabla f(x^k) + e^k\big),$$

$\Rightarrow \|u^{k+1}\| \le \|(I-sQ)u^k\| + s\|e^k\| \le q\|u^k\| + s\delta$

where s is a constant stepsize, e^k is an error satisfying $\|e^k\| \le \delta$ for all k, and f is the positive definite quadratic function

$$f(x) = \tfrac{1}{2}(x - x^*)'Q(x - x^*).$$

$\therefore \|u^k\| \le q\|u^{k-1}\| + s\delta$

$\le q(q\|u^{k-2}\| + s\delta) + s\delta$

Let

$$q = \max\big\{|1 - sm|, |1 - sM|\big\},$$

where

$\le \frac{s\delta}{1-q} + q^k\|u^0\|$

$$m: \text{ smallest eigenvalue of } Q, \qquad M: \text{ largest eigenvalue of } Q,$$

and assume that $q < 1$. Show that for all k, we have

Thus

$$\|x^k - x^*\| \le \frac{s\delta}{1 - q} + q^k\|x^0 - x^*\|.$$

(换元) let $u = x^k - x^*$

常数

Generated by CamScanner

to some vector e^* linearly, and in fact we have ... that $\{e^k\}$ converges

$$\|e^k - e^*\| \le q\beta^k$$

for some scalar q and all k. *Hint*: Show that $\{e^k\}$ is a Cauchy sequence.

1.3.11 (Convergence Rate of Steepest Descent with the Armijo Rule) (www)

P19 1.1.9

Let $f : \Re^n \mapsto \Re$ be a twice continuously differentiable function that satisfies

$$m\|y\|^2 \le y'\nabla^2 f(x)y \le M\|y\|^2, \qquad \forall\, x,\, y \in \Re^n,$$

$d^k = -\nabla f(x^k)$

where m and M are some positive scalars. Consider the steepest descent method $x^{k+1} = x^k - \alpha^k \nabla f(x^k)$ with α^k determined by the Armijo rule. Let x^* be the unique unconstrained minimum of f and let

$d^k = \beta^{m_k} s$ PPT 1 P43.

$$r = 1 - \frac{4m\beta\sigma(1-\sigma)}{M}.$$

(1) (2)

(2) 2 to

Show that for all k, we have

$$f(x^{k+1}) - f(x^*) \le r\big(f(x^k) - f(x^*)\big),$$

(3)

and

$$\|x^k - x^*\|^2 \le qr^k,$$

where q is some constant.

(4) → (5) ?

Solution

Generated by CamScanner

1.4.3 (Combination with Steepest Descent)

Consider the iteration $x^{k+1} = x^k + \alpha^k d^k$ where α^k is chosen by the Armijo rule with initial stepsize $s = 1$, $\sigma \in (0, 1/2)$, and d^k is equal to

$$d_N^k = -\left(\nabla^2 f(x^k)\right)^{-1} \nabla f(x^k)$$

if $\nabla^2 f(x^k)$ is nonsingular and the following two inequalities hold:

$$c_1 \left\|\nabla f(x^k)\right\|^{p_1} \leq -\nabla f(x^k)' d_N^k,$$

$$\|d_N^k\|^{p_2} \leq c_2 \left\|\nabla f(x^k)\right\|;$$

otherwise

$$d^k = -D\nabla f(x^k),$$

P35 where D is a fixed positive definite symmetric matrix. The scalars c_1, c_2, p_1, and p_2 satisfy $c_1 > 0$, $c_2 > 0$, $p_1 > 2$, $p_2 > 1$. Show that the sequence $\{d^k\}$ is gradient related. Furthermore, every limit point of $\{x^k\}$ is stationary, and if $\{x^k\}$ converges to a nonsingular local minimum x^*, the rate of convergence of $\{\|x^k - x^*\|\}$ is superlinear.

M cannot do it.

1.4.4 (Armijo Rule Along a Curved Path)

This exercise provides a globally

Generated by CamScanner