

Problem 1: Proving what Newton method is up to

Problem Statement

Newton's method minimizes the local quadratic approximation of a function, if its Hessian at the current iterate is positive definite. We showed that Newton's method indeed finds a *local* minimizer of the quadratic approximation. Here, we show that it finds the *global* minimizer through the following steps.

a

If Q is any PD matrix, show that:

$$\frac{1}{2}x^T Qx + c^T x = \frac{1}{2}(x + Q^{-1}c)^T Q(x + Q^{-1}c) - \frac{1}{2}c^T Q^{-1}c$$

b

Using the prior result, argue that the function $\frac{1}{2}x^T Qx + c^T x$ is minimized *globally* at $x^* = -Q^{-1}c$.

c

Recall that for any function $f : \mathbb{R} \rightarrow \mathbb{R}$, its local quadratic approximation at x^k is given by:

$$f^q(x) := f(x^k) + [\nabla f(x^k)]^T (x - x^k) + \frac{1}{2}(x - x^k)^T H(x^k)(x - x^k)$$

Assume the Hessian $H(x^k)$ is positive definite. Utilize the prior result to conclude that the *global* minimizer of f^q is given by $x^k - [H(x^k)]^{-1} \nabla f(x^k)$. Notice that the minimizer is indeed x^{k+1} , as defined by Newton's method.

Solution

a

b

c

Problem 2: Newton's method needs a touch-up

Problem Statement

In Newton's method, if the Hessian at the current iterate $H(x^k)$ is not PD, then $-[H(x^k)]^{-1} \nabla f(x^k)$ may not be a descent direction. Then, we modify the Hessian to $H(x^k) + D^k$ where D^k is a diagonal matrix with nonnegative diagonal entries. Let us design D^k to ensure that $H(x^k) + D^k$ is PD. The following corollary of Gershgorin's circle theorem will prove useful.

Theorem 0.1. *If λ is any eigenvalue of an arbitrary matrix $A \in \mathbb{R}^{n \times n}$, then*

$$|\lambda - A_{ii}| \leq \sum_{j \neq i} |A_{ij}|$$

for some $i = 1, \dots, n$.

a

Using this theorem, find a sufficient condition on the diagonal entries of A s.t. all eigenvalues of A are positive.

b

Using the prior result, find a diagonal matrix D^k s.t. that all eigenvalues of $H(x^k) + D^k$ are nonnegative, and all diagonal entries of D^k are nonnegative.

c

If any eigenvalue of $H(k) + D^k$ is close to zero but positive, then it is close to being singular and its inverse is susceptible to noise. Modify your answer in the prior section to ensure that all the eigenvalues are greater than $\frac{1}{2}$.

d

The file `applyNewtonMethod.m` implements a basic newton method to the function

$$f(x_1, x_2) := \cos(x_1^2 - 2x_2) + \sin(x_1^2 + x_2^2)$$

starting from $x^0 = (1.2, 0.5)$. The program also draws the contour plot and the surface plot of f .

1. Verify the Hessian at the starting point is NOT PD.
2. Does the Newton method converge? If yes, does it converge to a local minimizer of f ?
3. Fill in the missing code in `modifyHessian.m` that takes $H(x^k)$ as input and gives $H(x^k) + D^k$ as output. Utilize your condition in part (c) to design D^k . Submit your code. Using your code, compute $H(x^0) + D^0$; i.e the modified Hessian at the starting point.
4. Uncomment the relevant lines in `applyNewtonMethod.m` to run the modified Newton method. Report if the algorithm converges to a local minimizer of f .

Solution

a