Newton step: $\boxed{\Delta x_{nt} = -\left(\nabla^2 f(x)\right)^{-1} \nabla f(x).}$

If $\nabla^2 f(x) > 0$,

$$\nabla f(x)^T \cdot \Delta x_{nt} = -\nabla f(x)^T \left(\nabla^2 f(x)^{-1}\right) \nabla f(x) < 0.$$

$x^T A x > 0 \quad \forall x \neq 0$ if $A > 0$.

unless $\nabla f(x) = 0$. or x is optimal.

Thus the Newton step is a decent direction.

— Δx_{nt} minimizes the second-order Taylor approximation of f at x:

$$\hat{f}(x+v) = f(x) + \nabla f(x)^T v + \frac{1}{2} v^T \nabla^2 f(x) v.$$

$0 = \frac{\partial \hat{f}(x+v)}{\partial v} \Rightarrow v^* = \Delta x_{nt}$. (Check this).

— If f is quadratic, $\rightarrow \hat{f} = f$ then

$x + \Delta x_{nt}$ is the exact minimizer of f.

— For f twice differentiable, and f is close to quadratic the quadratic model of f will very accurate when x is near x^*.

— $x + \Delta x_{nt}$ is a good estimate of x^*.

The Newton decrement

$$\lambda(x) = \left(\nabla f(x)^T \underbrace{\nabla^2 f(x)^{-1} \nabla f(x)}_{-\nabla f(x)^T \cdot \Delta x_{nt}}\right)^{1/2}$$

- The Newton decrement is an estimate of $f(x) - p^*$ based on the quadratic approximation $\hat{f}$ at x.

$$f(x) - \inf_y \hat{f}(y) = f(x) - \hat{f}(x + \Delta x_{nt}) = \frac{1}{2}\lambda(x)^2.$$

- $\lambda(x) = \left(\Delta x_{nt}^T \nabla^2 f(x) \Delta x_{nt}\right)^{1/2}.$

$\Delta x_{nt} = -\nabla^2 f(x)^{-1} \nabla f(x).$

Newton's method

Given a starting point $x \in \operatorname{dom} f$.
and $\epsilon > 0$.

Repeat 1. Compute the Newton step

$$\Delta x_{nt} := -\left(\nabla^2 f(x)\right)^{-1} \nabla f(x)$$

$f(x) = \log(e^{x} + e^{-x})$
$x \in \mathbb{R}$.

and the Newton decrement

$$\lambda^2 := \nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x).$$

$f'(x)$.
$f''(x)$.

2. Stopping criterion. quit if $\lambda^2/2 \leq \epsilon$.

3. Line search. Choose step size t by backtracking line search.

4. Update $x := x + t\Delta x_{nt}$.

If $t=1$ is fixed, it is call the pure Newton method.. (no line search).
Otherwise it is called the damped Newton method.

Equality constrained minimization problems.

$$\text{minimize } f(x)$$
$$\text{subject to } Ax=b.$$

→ dual problem is unconstrained.

where $f(x): \mathbb{R}^n \to \mathbb{R}$ is convex and twice continuously differentiable and $A \in \mathbb{R}^{p\times n}$ with $\text{rank}(A) = p < n$.

- Optimality condition (KKT condition)

x^* is optimal if there exists a ν^* such that

$$\boxed{\nabla f(x^*) + A^T \nu^* = 0, \quad Ax^* = b.}$$

feasibility of the primal problem

$$L(x, \nu) = f(x) + \nu^T (Ax - b).$$

$$0 = \nabla_x L(x^*, \nu^*)$$

§10.2

Newton's method with equality constraints.

- The initial point must be feasible.
  (i.e. $x^{(0)} \in \operatorname{dom} f$, and $Ax^{(0)} = b$).

- The definition of the Newton step is modified to take the equality constraints into account. The Newton step Δx_{nt} is a feasible direction, i.e. $\boxed{A \Delta x_{nt} = 0}$.

$$x^+ := x + t \Delta x_{nt}.$$

Need $\underline{Ax^+ = b}$.

$$A(x + t\Delta x_{nt}) = Ax + t\underbrace{(A \Delta x_{nt})}_{0}$$

$$= b$$

Δx_{nt} is determined by solving the second-order approximation (with variable v)

$$\text{minimize} \quad \hat{f}(x+v) = f(x) + \nabla f(x)^T v + \frac{1}{2} v^T \nabla^2 f(x) v$$

$$\text{s.t.} \quad A(x+v) = b$$

$$\boxed{\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix}} \begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}.$$

called KKT matrix.

$$\begin{cases} \nabla^2 f(x) v + A^T w = -\nabla f(x). \\ A v = 0. \end{cases}$$