CME 307: Optimization	CME 307 / MS&E 311 / OIT 676
Lecture 12: Newton and quasi-Newton	
Fall 2025	*Prof. Udell*

1 Newton and quasi-Newton methods

We study algorithms for smooth unconstrained minimization

$$\min_{x \in \mathbb{R}^n} f(x),$$

with f continuously differentiable and, when needed, twice continuously differentiable. We will (i) build and minimize local quadratic models, (ii) derive Newton's method and its local quadratic convergence, (iii) introduce quasi-Newton updates (BFGS and L-BFGS), and (iv) discuss globalization (line search and trust regions), preconditioning, and variable-metric interpretations. Throughout we recall basic notions such as L-smoothness and strong convexity as needed.

1.1 Quadratic models and search directions

Definition 1.1 (Quadratic model)**.** Given $x_k \in \mathbb{R}^n$ and a symmetric matrix B_k, the second-order (quadratic) model of f at x_k is

$$m_k(s) = f(x_k) + \nabla f(x_k)^\top s + \tfrac{1}{2} s^\top B_k s, \qquad s \in \mathbb{R}^n.$$

The associated trial point is $x_{k+1} = x_k + s_k$ where s_k approximately minimizes m_k. If B_k is positive definite (pd), the unique minimizer is $s_k = -B_k^{-1} \nabla f(x_k)$.

Remark 1.2 (Why $B_k \succ 0$?). If B_k is indefinite, m_k is unbounded below along directions of negative curvature; if $B_k \succeq 0$ but singular, the minimizer may not be unique or may not reduce f sufficiently along flat directions. Ensuring $B_k \succ 0$ yields a unique descent direction s_k.

1.2 Globalization: line search and trust regions

We will combine directions from a quadratic model with a globalization strategy to guarantee global convergence. The most common globalization strategy is linesearch, which we describe below.

Backtracking linesearch and Armijo condition. Given a descent direction p_k (i.e., $\nabla f(x_k)^\top p_k < 0$), backtracking linesearch chooses the largest $\alpha \in 1, \beta, \beta^2, \ldots$ that satisfies the *Armijo condition*,

$$f(x_k + \alpha p_k) \le f(x_k) + c_1 \alpha, \nabla f(x_k)^\top p_k,$$

with fixed $c_1 \in (0,1)$ and $\beta \in (0,1)$.