THE LOCAL ANALYSIS APPROACH

- Restrict attention to sequences x^k converging to a local min x^*
- Measure progress in terms of an error function $e(x)$ with $e(x^*) = 0$, such as

$$e(x) = \|x - x^*\|, \qquad e(x) = f(x) - f(x^*)$$

- Compare the tail of the sequence $e(x^k)$ with the tail of standard sequences
- Geometric or linear convergence [if $e(x^k) \leq q\beta^k$ for some $q > 0$ and $\beta \in [0, 1)$, and for all k]. Holds if

$$\limsup_{k \to \infty} \frac{e(x^{k+1})}{e(x^k)} < \beta$$

- Superlinear convergence [if $e(x^k) \leq q \cdot \beta^{p^k}$ for some $q > 0$, $p > 1$ and $\beta \in [0, 1)$, and for all k].
- Sublinear convergence

QUADRATIC MODEL ANALYSIS

- Focus on the quadratic function $f(x) = (1/2)x'Qx$, with $Q > 0$.
- Analysis also applies to nonquadratic problems in the neighborhood of a nonsingular local min
- Consider steepest descent

$$x^{k+1} = x^k - \alpha^k \nabla f(x^k) = (I - \alpha^k Q)x^k$$

$$\begin{aligned}\|x^{k+1}\|^2 &= x^{k'}(I - \alpha^k Q)^2 x^k \\ &\leq \big(\text{max eig. } (I - \alpha^k Q)^2\big)\|x^k\|^2\end{aligned}$$

The eigenvalues of $(I - \alpha^k Q)^2$ are equal to $(1 - \alpha^k \lambda_i)^2$, where λ_i are the eigenvalues of Q, so

$$\text{max eig of } (I-\alpha^k Q)^2 = \max\big\{(1-\alpha^k m)^2, (1-\alpha^k M)^2\big\}$$

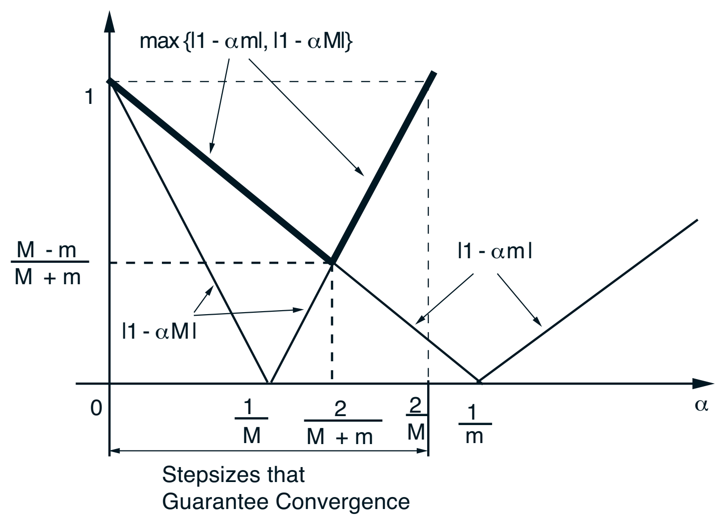
where m, M are the smallest and largest eigenvalues of Q. Thus

$$\frac{\|x^{k+1}\|}{\|x^k\|} \leq \max\big\{|1 - \alpha^k m|, |1 - \alpha^k M|\big\}$$

OPTIMAL CONVERGENCE RATE

- The value of α^k that minimizes the bound is $\alpha^* = 2/(M+m)$, in which case

$$\frac{\|x^{k+1}\|}{\|x^k\|} \le \frac{M-m}{M+m}$$



- Conv. rate for minimization stepsize (see text)

$$\frac{f(x^{k+1})}{f(x^k)} \le \left(\frac{M-m}{M+m}\right)^2$$

- The ratio M/m is called the *condition number* of Q, and problems with M/m: large are called *ill-conditioned*.