6.252 NONLINEAR PROGRAMMING

LECTURE 3: GRADIENT METHODS

LECTURE OUTLINE

- Quadratic Unconstrained Problems
- Existence of Optimal Solutions
- Iterative Computational Methods
- Gradient Methods - Motivation
- Principal Gradient Methods
- Gradient Methods - Choices of Direction

QUADRATIC UNCONSTRAINED PROBLEMS

$$\min_{x \in \Re^n} f(x) = \tfrac{1}{2} x' Q x - b' x,$$

where Q is $n \times n$ symmetric, and $b \in \Re^n$.

- Necessary conditions:

$$\nabla f(x^*) = Q x^* - b = 0,$$

$$\nabla^2 f(x^*) = Q \geq 0 : \text{ positive semidefinite.}$$

- $Q \geq 0 \;\Rightarrow\; f$: convex, nec. conditions are also sufficient, and local minima are also global

- Conclusions:
  - Q : not $\geq 0 \;\Rightarrow\; f$ has no local minima
  - If $Q > 0$ (and hence invertible), $x^* = Q^{-1} b$ is the unique global minimum.
  - If $Q \geq 0$ but not invertible, either no solution or ∞ number of solutions

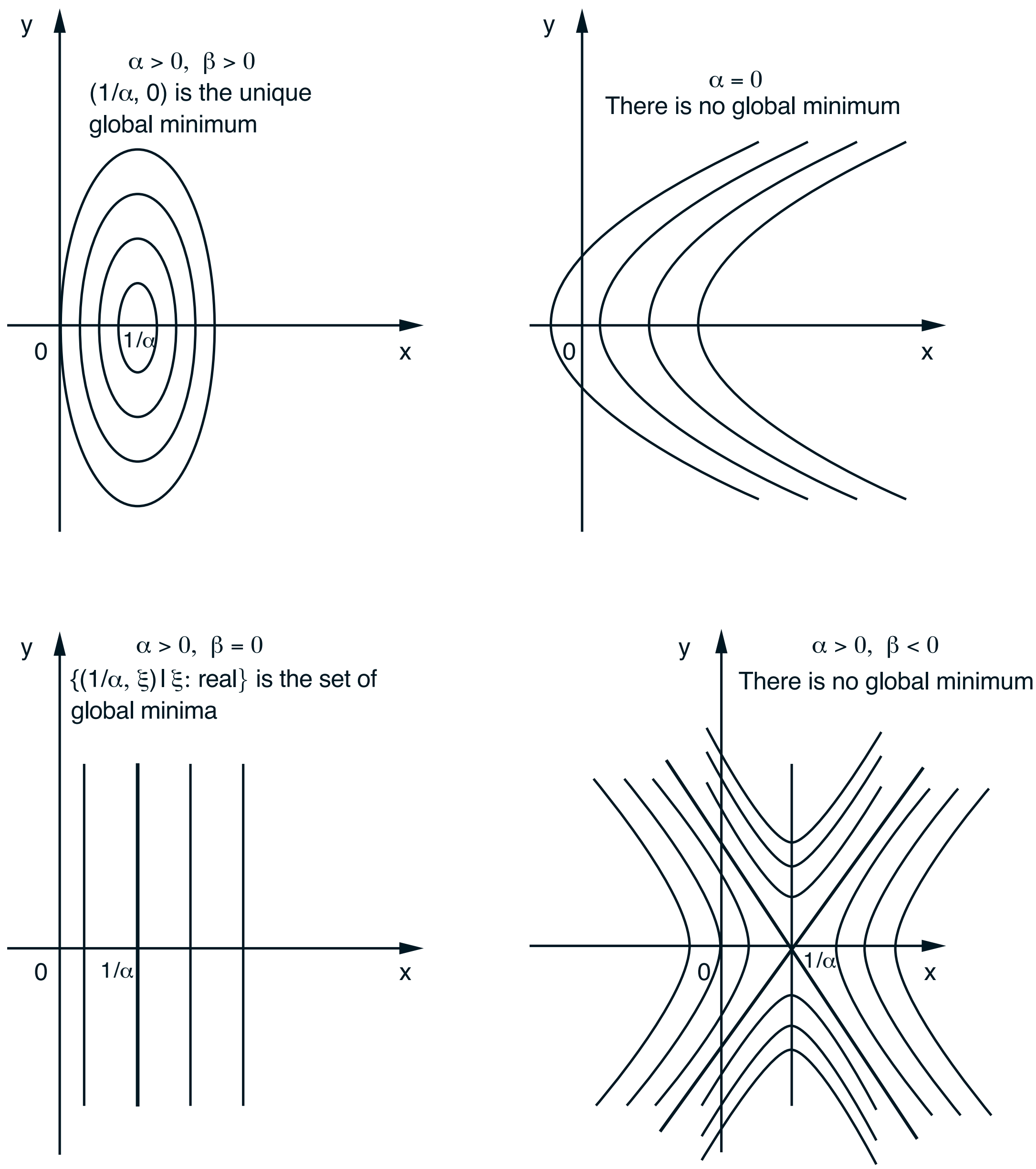


Illustration of the isocost surfaces of the quadratic cost function $f : \Re^2 \mapsto \Re$ given by

$$f(x, y) = \tfrac{1}{2}\left(\alpha x^2 + \beta y^2\right) - x$$

for various values of α and β.

EXISTENCE OF OPTIMAL SOLUTIONS

Consider the problem

$$\min_{x \in X} f(x)$$

- The set of optimal solutions is

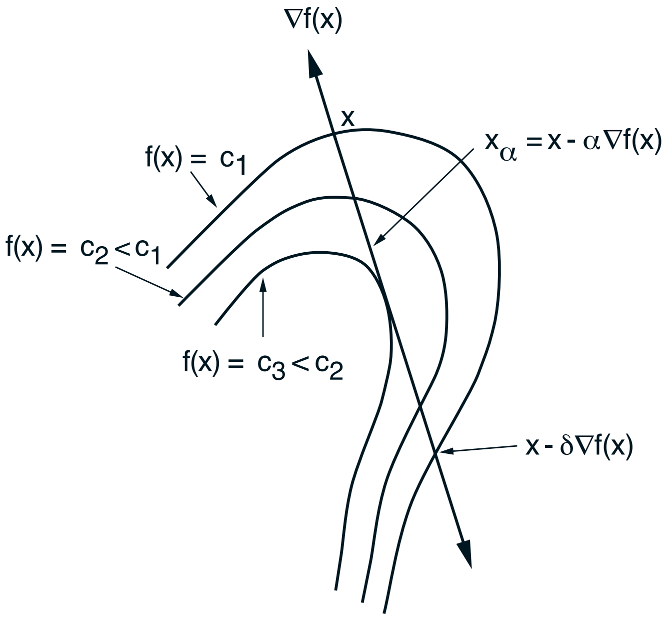
$$X^* = \cap_{k=1}^{\infty} \{x \in X \mid f(x) \leq \gamma_k\}$$

where $\{\gamma_k\}$ is a scalar sequence such that $\gamma_k \downarrow f^*$ with

$$f^* = \inf_{x \in X} f(x)$$

- X^* is nonempty and compact if all the sets $\{x \in X \mid f(x) \leq \gamma_k\}$ are compact. So:
  - A global minimum exists if f is continuous and X is compact (Weierstrass theorem)
  - A global minimum exists if X is closed, and f is continuous and coercive, that is, $f(x) \to \infty$ when $\|x\| \to \infty$

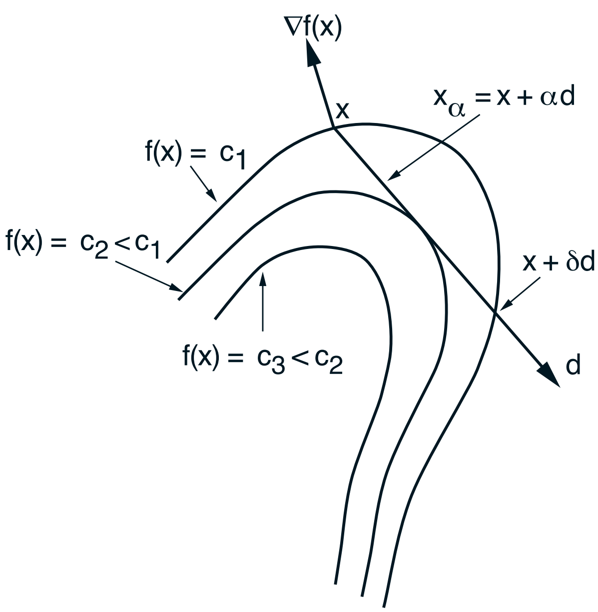
GRADIENT METHODS - MOTIVATION



If $\nabla f(x) \neq 0$, there is an interval $(0, \delta)$ of stepsizes such that

$$f\Big(x - \alpha\nabla f(x)\Big) < f(x)$$

for all $\alpha \in (0, \delta)$.



If d makes an angle with $\nabla f(x)$ that is greater than 90 degrees,

$$\nabla f(x)'d < 0,$$

there is an interval $(0, \delta)$ of stepsizes such that $f(x+\alpha d) < f(x)$ for all $\alpha \in (0, \delta)$.

PRINCIPAL GRADIENT METHODS

$$x^{k+1} = x^k + \alpha^k d^k, \qquad k = 0, 1, \ldots$$

where, if $\nabla f(x^k) \neq 0$, the direction d^k satisfies

$$\nabla f(x^k)' d^k < 0,$$

and α^k is a positive stepsize. Principal example:

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

where D^k is a positive definite symmetric matrix

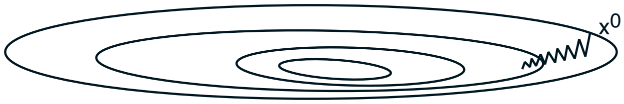
- Simplest method: **Steepest descent**

$$x^{k+1} = x^k - \alpha^k \nabla f(x^k), \qquad k = 0, 1, \ldots$$

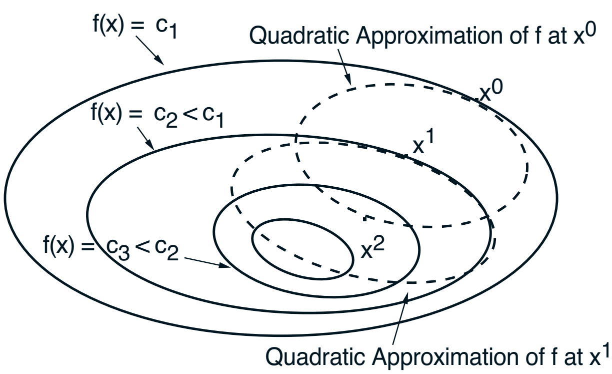
- Most sophisticated method: **Newton's method**

$$x^{k+1} = x^k - \alpha^k \big(\nabla^2 f(x^k)\big)^{-1} \nabla f(x^k), \qquad k = 0, 1, \ldots$$

STEEPEST DESCENT AND NEWTON'S METHOD



Slow convergence of steepest descent



Fast convergence of Newton's method w/ $\alpha^k = 1$.

Given x^k, the method obtains x^{k+1} as the minimum of a quadratic approximation of f based on a second order Taylor expansion around x^k.

OTHER CHOICES OF DIRECTION

- **Diagonally Scaled Steepest Descent**

$$D^k = \text{ Diagonal approximation to } \big(\nabla^2 f(x^k)\big)^{-1}$$

- **Modified Newton's Method**

$$D^k = (\nabla^2 f(x^0))^{-1}, \qquad k = 0, 1, \ldots,$$

- **Discretized Newton's Method**

$$D^k = \big(H(x^k)\big)^{-1}, \qquad k = 0, 1, \ldots,$$

where $H(x^k)$ is a finite-difference based approximation of $\nabla^2 f(x^k)$,

- **Gauss-Newton method** for least squares problems: $\min_{x \in \Re^n} \frac{1}{2}\|g(x)\|^2$. Here

$$D^k = \big(\nabla g(x^k) \nabla g(x^k)'\big)^{-1}, \qquad k = 0, 1, \ldots$$