Descent methods

$$x^{(k+1)} = x^{(k)} + t^{(k)} \Delta x^{(k)} \quad \text{with } f(x^{(k+1)}) < f(x^{(k)})$$

- other notations: $x^+ = x + t\Delta x$, $x := x + t\Delta x$

- Δx is the *step*, or *search direction*; t is the *step size*, or *step length*

- from convexity, $f(x^+) < f(x)$ implies $\nabla f(x)^T \Delta x < 0$
  (*i.e.*, Δx is a *descent direction*)

General descent method.

given a starting point $x \in \mathbf{dom}\, f$.
repeat
      1. Determine a descent direction Δx.
      2. *Line search.* Choose a step size $t > 0$.
      3. *Update.* $x := x + t\Delta x$.
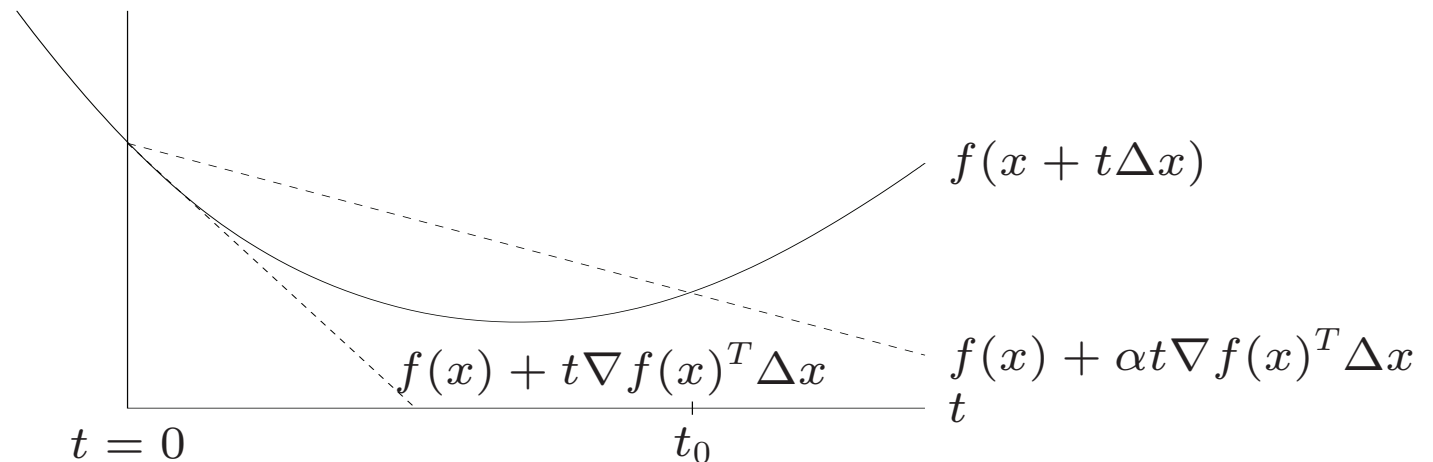**until** stopping criterion is satisfied.

Line search types

exact line search: $t = \mathrm{argmin}_{t>0} \, f(x + t\Delta x)$

backtracking line search (with parameters $\alpha \in (0, 1/2)$, $\beta \in (0, 1)$)

- starting at $t = 1$, repeat $t := \beta t$ until

$$f(x + t\Delta x) < f(x) + \alpha t \nabla f(x)^T \Delta x$$

- graphical interpretation: backtrack until $t \leq t_0$

Gradient descent method

general descent method with $\Delta x = -\nabla f(x)$

---

given a starting point $x \in \mathbf{dom}\, f$.
repeat
    1. $\Delta x := -\nabla f(x)$.
    2. *Line search.* Choose step size t via exact or backtracking line search.
    3. *Update.* $x := x + t\Delta x$.
until stopping criterion is satisfied.

- stopping criterion usually of the form $\|\nabla f(x)\|_2 \le \epsilon$

- convergence result: for strongly convex f,

$$f(x^{(k)}) - p^\star \le c^k(f(x^{(0)}) - p^\star)$$

  $c \in (0, 1)$ depends on m, $x^{(0)}$, line search type

- very simple, but often very slow; rarely used in practice

quadratic problem in $\mathbf{R}^2$

$$f(x) = (1/2)(x_1^2 + \gamma x_2^2) \qquad (\gamma > 0)$$

with exact line search, starting at $x^{(0)} = (\gamma, 1)$:

$$x_1^{(k)} = \gamma \left(\frac{\gamma - 1}{\gamma + 1} \right)^k, \qquad x_2^{(k)} = \left(-\frac{\gamma - 1}{\gamma + 1} \right)^k$$

- very slow if $\gamma \gg 1$ or $\gamma \ll 1$

- example for $\gamma = 10$: