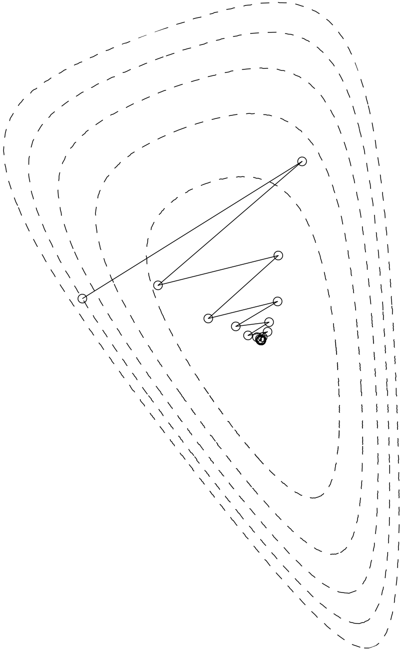

Optimization Algorithms

Overview – Downhill Algorithms for Unconstrained Optimization

Marc Toussaint
Technical University of Berlin
Winter 2024/25

Problem Formulation

- Unconstrained non-linear mathematical program:

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

  - for smooth function $f : \mathbb{R}^n \to \mathbb{R}$
  - we can query $f(x)$, $\nabla f(x)$ (gradient methods), and sometimes $\nabla^2 f(x)$ (2nd order methods)

Problem Formulation

- Unconstrained non-linear mathematical program:

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

  - for smooth function $f : \mathbb{R}^n \to \mathbb{R}$
  - we can query $f(x)$, $\nabla f(x)$ (gradient methods), and sometimes $\nabla^2 f(x)$ (2nd order methods)

- Application examples in robotics, model fitting, parameter optimization, etc.

We aim for methods that are:

- **monotone**
  - generates a sequence of points x_i that gradually reduce the value $f(x_i) \leq f(x_{i\text{-}1})$
- **convergent**
  - under bounded positive curvature assumptions we want exponential convergence rates guaranteed
- **invariant to rescaling of** f
- **invariant to rescaling of** x (or to linear transform of x)

Note on local vs. global optimization

- Part I and II focus on local optimization – the key challenge is to converge to local optima as fast as possible

- Global optimization requires to search for various local optima
  - Restart local downhill solvers from various points
  - Use Bayesian Optimization or other explicit global search concepts
    $\rightarrow$ Global Optimization lecture

Note on convex vs. non-convex optimization

- The methods we discuss equally apply to convex and non-convex problems
  - If convex (with bounded curvature) they have strong convergence guarantee
  - If non-convex, they still run downhill to local minima, but without the same guarantee

Outline

- Gradient descent, **stepsize** adaptation, & backtracking line search

- Steepest descent **direction**, Newton, damping & non-convex fallback, trust region
- Quasi-Newton, Gauss-Newton, BFGS, conjugate gradient