A QUICK TOUR ON NUMERICAL OPTIMIZATION

Seungjin Lee
ETRI

24. 07. 2023 and 27. 07. 2023

Part I

optimization problems and optimization algorithms

introduction

optimization problems

optimization algorithms

constrained and unconstrained problems

unconstrained optimization problems

(unconstrained) optimization algorithms

why optimization matters?

> *People optimize.*
>
> *Numerical Optimization*, Nocedal and Wright

introduction

optimization problems

optimization algorithms

constrained and unconstrained problems

unconstrained optimization problems

(unconstrained) optimization algorithms

optimization problem: abstract construction

An optimization problem consists of the followings:

- a formula $\varphi(x)$,
- a structure $\mathcal{U}$ providing a logical semantic of $\varphi(x)$ for $x \in |\mathcal{U}|$.

optimization problem: practical construction

In practice, we have

- a set X of *variables*,
- an ordered set Y,
- an *objective function* $X \xrightarrow{f} Y$,
- the formula $\varphi(x) = \sup_{x \in X} f(x)$.

optimization problem: classification

An optimization problem $(\mathcal{U}, \varphi(x))$ is classified by

- the structure $\mathcal{U}$,
- the formula $\varphi(x)$.

optimization problem: classification

For an optimization problem $(X, Y, X \xrightarrow{f} Y)$

- X is explicitly given $\rightarrow$ unconstrained problem,
- X is implicitly given by *constraints* $\rightarrow$ constrained problem,
- X is a discrete set $\rightarrow$ discrete problem,
- X is a metric space $\rightarrow$ continuous problem,
- X is a probability space $\rightarrow$ stochastic optimization.

optimization problem: classification

For an optimization problem $(X, Y, X \xrightarrow{f} Y)$:

- f is linear $\rightarrow$ linear problem,
- f is convex $\rightarrow$ convex problem,
- f is nonlinear $\rightarrow$ nonlinear problem,
- f is differentiable $\rightarrow$ differentiable problem, and
- ...(quantum?).

introduction

optimization problems

optimization algorithms

constrained and unconstrained problems

unconstrained optimization problems

(unconstrained) optimization algorithms

optimization algorithm

Given an optimization problem $(\varphi(x), \mathcal{U})$, an *optimization algorithm* is a Turing machine $\mathcal{T}_{\text{opt}}$ which accepts sequences on X having the last element or converging to x^* satisfying $\varphi(x^*)$.

optimization algorithm: classification

For an optimization algorithm $\mathcal{T}_{\text{opt}}$ of $(X, Y, X \xrightarrow{f} Y)$:

- $\sup_X f(x^*) \to$ global optimization,
- $\sup_{U \subset X} f(x^*) \to$ local optimization.

introduction

optimization problems

optimization algorithms

constrained and unconstrained problems

unconstrained optimization problems

(unconstrained) optimization algorithms

constrained optimization problem

A constrained optimization is casted into the form of

$$\min_{x \in X} f(x) \quad \text{subject to} \quad \begin{cases} c_E(x) = 0, \\ c_I(x) \geq 0. \end{cases}$$

constrained optimization problem: an illusion?

$$\min_{x \in \mathbb{R}} f(x) \quad \text{subject to} \quad \begin{cases} x + 1 \geq 0, \\ 1 - x \geq 0 \end{cases}$$

vs

$$\min_{\theta \in \mathbb{R}} f(\cos \theta)$$

constrained system

Relativistic point-particle:

$$(x^a, p_a) \quad \text{subject to} \quad p_a p^a + m^2 = 0.$$

For a massless particle, the mass-shell condition can be solved by twistors $\omega, \pi \in \mathbb{C}^2$:

$$p_a p^a = 0 \leftrightarrow \begin{cases} \omega^A = i x^a \gamma_a^{AB'} \pi_{B'}, \\ \bar{\omega}^{B'} = -i x^a \gamma_a^{AB'} \bar{\pi}_A, \\ p_a \gamma^a_{AB'} = \pi_A \bar{\pi}_{B'} \end{cases}$$

introduction

optimization problems

optimization algorithms

constrained and unconstrained problems

unconstrained optimization problems

(unconstrained) optimization algorithms

unconstrained optimization problem

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

- x^* is a *global minimizer* if $f(x^*) \leq f(x), \ \forall x \in X$.
- x^* is a *local minimizer* if there exists an open set $U : x^* \in U$ such that $f(x^*) \leq f(x^*), \ \forall x \in U$,
- x^* is a *strict local minimizer* if there exists an open set $U : x^* \in U$ such that $f(x^*) < f(x), \ \forall (x \in U \wedge x \neq x^*)$.

NOTATION

$$\partial_a = \frac{\partial}{\partial x^a}, \quad X^{ab\dots} Y_{ac\dots} = \sum_a X^{ab\dots} Y_{ac\dots}$$

unconstrained optimization problem: local solution

For a continuous differentiable $\mathbb{R}^n \xrightarrow{f} \mathbb{R}$

$$x^* \text{ is a local minimizer} \quad \rightarrow \quad \begin{gathered} \partial_a f(x^*) = 0 \\ p^a p^b \partial_a \partial_b f(x^*) \geq 0 \end{gathered}$$

$$x^* \text{ is a strict local solution} \quad \leftarrow \quad \begin{gathered} \partial_a f(x^*) = 0 \\ p^a p^b \partial_a \partial_b f(x^*) > 0 \end{gathered}$$

convex optimization problem

A *convex set* is a set C endowed with the addition such that for any $x_1, x_2 \in C$ and $0 \leq \theta \leq 1$

$$\theta x_1 + (1 - \theta)x_2 \in C.$$

A *convex function* is a function $C \xrightarrow{f} Y$ such that

$$f(\theta x_1 + (1 - \theta)x_2) \leq \theta f(x_1) + (1 - \theta)f(x_2).$$

convex optimization problem: good news

For a convex optimization problem,

- any local minimizer is a global minimizer,
- if $\partial_a f(x^*) = 0$, then x^* is a global minimizer.

introduction

optimization problems

optimization algorithms

constrained and unconstrained problems

unconstrained optimization problems

(unconstrained) optimization algorithms

line search and trust region methods

line search method	$\min_{\alpha>0} f(x_k + \alpha p_k)$
	step-length α
	search direction p_k
trust region method	$\min_p \hat{f}(x_k + p)$
	model function $\hat{f}(x + p)$
	step p

line search method: steepest descent

For $0 < \alpha \ll 1$,

$$f(x_k + \alpha p) \approx f(x_k) + \alpha p^a \frac{\partial}{\partial x^a} f(x_k)$$

Maximally decreasing if

$$p_a = -\frac{\partial}{\partial x^a} f(x_k) \left\| \frac{\partial}{\partial x^b} f(x_k) \right\|^{-1}$$

line search method: Newton direction

Minimize the following:

$$f(x_k + p) \approx f(x_k) + p^a \frac{\partial}{\partial x^a} f(x_k) + \frac{1}{2} p^a p^b \frac{\partial^2}{\partial x^a \partial x^b} f(x_k)$$

The Newton direction

$$p_a^k = -\left(\frac{\partial^2}{\partial x^a \partial x^b} f(x_k) \right)^{-1} \frac{\partial}{\partial x^a} f(x_k)$$

line search method: Newton direction

Require

$$0 = \partial_a f(x_k + p) \approx \frac{\partial}{\partial x^a} f(x_k) + p^b \frac{\partial^2}{\partial x^a \partial x^b} f(x_k)$$

while $f(x_k) \geq f(x_k + p)$.

line search method: BFGS formula

Observe that near a solution

$$\frac{\partial}{\partial x^a} f(x_{k+1}) - \frac{\partial}{\partial x^a} f(x_k) \approx \frac{\partial^2}{\partial x^a \partial x^b} f(x_k)(x_{k-1} - x_k)^b.$$

Let

$$s_k = (x_{k+1} - x_k), \quad y_k = \frac{\partial}{\partial x^a} f(x_{k+1}) - \frac{\partial}{\partial x^a} f(x_k).$$

Approximate the Hessian by

$$(B_{k+1})_{ab} = (B_k)_{ab} + \frac{(y_k - B_k s_k)_a (y_k - B_k s_k)_b}{(y_k - B_k s_k)_c s_k^c}.$$

(the Broyden-Fletcher-Goldfarb-Shanno formula)