Non-Deterministic Local Search Methods

*21871-01 Optimization*

Dietmar Maringer

WWZ, University of Basel

Spring 2012
The Problem with Optimization Problems

problem space and complexity

– discontinuous
– non-convex
– constrained
The Problem with Optimization Problems

methods (revisited)
  – closed-form solution | exact algorithm
  – approximations
    • simplified problem
    • rule of the thumb
    • Monte Carlo methods
  – numerically
    • methods based on FOC
    • heuristics
Non-Deterministic Search Methods

stochastic and heuristic methods
– use random elements for
  • creating new candidate solutions
  • accepting new solutions
– non-deterministic
– often: inspired by nature
– flexible
Stochastic Differential Equations (SDE)

basic idea

– guided search: gradient $\nabla f$

– escape local optima: noise $\epsilon_i, E(\epsilon_i) = 0$

“cooling sequence” $T(i)$

– generation of new candidate solutions

$$x_i = x_{i-1} \pm \lambda \cdot \nabla f + \epsilon_i \cdot T(i)$$
Stochastic Differential Equations (SDE)

application: portfolio optimization (Maringer & Parpas (2009))

- objective
\[
\min_x f(x) = \lambda \mathbb{V}(r_P) - (1 - \lambda) \mathbb{E}(r_P)
\]

- constraints
\[
r_{p,t} = \sum_i x_i r_{i,t} \quad x_i \geq 0 \quad \forall i \quad \sum_i x_i = 1
\]

- constraint on higher moment (skewness / kurtosis)
\[
\mathbb{M}(r_P) = \mathbb{M}^*
\]

- diffusion process and correction term
\[
dX(t) = -\nabla f(X(t)) dt + \sqrt{2T(t)} dB(t) - \nabla g(X(t))^T \lambda(X(t), t) dt
\]
Non-Deterministic Search Methods

gradient search + noise

– advantages
  • easy to implement
  • based on (numerical approximation of) 1\textsuperscript{st} derivative
  • can overcome local optima

– disadvantages
  • not suitable for “extremely rough” solution space
  • initial solution still (somewhat) important
  • convergence behaviour
  • computation of gradient can be computationally costly
Non-Deterministic Search Methods

annealing in the physical world

– goal
  • obtaining low energy states
  • particles are arranged in highly structured lattices

– possible movements of the particles:
  • decreasing the system’s energy: high probability
  • increasing the system’s energy: low probability
  • probabilities depend on temperature and magnitude of change in energy state

– stable and good terminal states only when
  • initial temperature is sufficiently high
  • cooling is sufficiently slow
  • terminal temperature is sufficiently low (freezing)
Simulated Annealing (SA)

– Kirkpatrick, Gelatt, Vecchi (Science, 1983)
– mimics the natural annealing process
  • improvements are (likely to be) accepted
  • impairments are (increasingly) likely to be rejected
  • probabilities depend on
    – “temperature” (i.e., progress)
    – magnitude of $\Delta f$
Simulated Annealing

maximisation of function $f(x)$ with SA

hill-climbing method

Find **good** initial solution, $x$;

REPEAT

make sophisticated guess for $\Delta x$;

$x_{new} := x + \Delta x$;

if $f(x_{new}) - f(x) \geq 0$

then $x := x_{new}$;

UNTIL converged

Simulated Annealing

(Kirkpatrick, Gellat, Vecchi, 1983)

Find **random** initial solution, $x$;

REPEAT

make random guess for $\Delta x$;

$x_{new} := x + \Delta x$;

with prob$(f(x_{new}) - f(x), \text{Temp.})$

$x := x_{new}$;

lower Temperature;

UNTIL halting criterion met;
Simulated Annealing

Accepting New Solutions in SA (max problem)

Candidate solution: A \(\rightarrow\) B

Current solution: A

\[ f(x) \]

\( \text{prob}(Df; T) \) = ACCEPT

\( \text{prob}(Df; T) \) = REJECT
Simulated Annealing

probability of accepting a new solution (max problem)

(a) SA with Metropolis function: \[ p = \begin{cases} 1 & \Delta f > 0 \\ \exp (\Delta f / T) & \Delta f < 0 \end{cases} \]

(b) SA with Boltzmann function: \[ p = 1 / (1 + \exp (-\Delta f / T)) \]

(c) Threshold Accepting: \[ p = \begin{cases} 1 & \Delta f > \tau \text{ where } \tau < 0 \\ 0 & \Delta f < \tau \end{cases} \]
Threshold Accepting

maximisation of function $f(x)$ with TA

**hill-climbing methods**

Find **good** initial solution, $x$;

REPEAT

make **sophisticated** guess for $\Delta x$;
$x^{new} := x + \Delta x$;

if $f(x^{new}) - f(x) \geq 0$
then $x := x^{new}$;

UNTIL converged

---

**Threshold Accepting**

(Dueck and Scheuer, 1990)

Find **random** initial solution, $x$;

REPEAT

make **random** guess for $\Delta x$;
$x^{new} := x + \Delta x$;

if $f(x^{new}) - f(x) \geq $ Threshold
then $x := x^{new}$;
reduce Threshold;

UNTIL **halting criterion met**;
Simulated Annealing and Threshold Accepting

implementation issues

– Simulated Annealing and Threshold Accepting
  • number of iterations
  • neighbourhood (i.e., distribution of $\Delta x$)
  • acceptance:
    – SA: initial temperature, cooling sequence, acceptance function
    – TA: initial threshold, threshold sequence

– constraint satisfaction
  • allow feasible solutions only (neighbourhood definition)
  • repair mechanism
    – generate new solution
    – if infeasible: find “nearest” feasible solution
  • punishment
    – impairment of objective function when constraints are violated
Simulated Annealing and Threshold Accepting

practical issues

– neighbourhoods for (multiple) decision variables
  • number of variables changed in one step
  • suitable changes with respect to
    – type (e.g., integer ⇔ real)
    – magnitude

– local updating
  • given suitable neighbourhood definition, \( \Delta f \) can be computed directly from \( \Delta x \) and \( f(x) \)
  • advantage: might be faster (CPU time!)
  • example
    – objective function: sum of (transformed) values of \( x_i \)
    – only few elements in the objective function are concerned (i.e., elements where \( x_i \) changed)
Simulated Annealing and Threshold Accepting

practical issues

– threshold sequence / cooling sequence

• distributions of $\Delta x$ and $\Delta f$ are related

• empirical approach:
  – assume distribution of $\Delta x$ for search process
  – estimate resulting distribution of $\Delta f$
  – choice initial temperature (threshold) and cooling (threshold) sequence:
    » high acceptance rate for impairments in the beginning
    » very high rejection rate for impairments towards the end
Simulated Annealing and Threshold Accepting

Literature


(older version: WP009-06)