Third LACM-EFES-JUSTIPEN Workshop
Joint Institute for Heavy Ion Research
Oak Ridge National Laboratory
February 23-25, 2009

Recent Developments in Parameter Estimation for Complex Simulations

Jorge Moré

Joint work with Jason Sarich and Stefan Wild

Mathematics and Computer Science Division
Argonne National Laboratory

Argonne NATIONAL LABORATORY

Complex Simulations

Parameter estimation problem are of the form

$$\min \{ f(x) : x_L \le x \le x_U \}, \qquad f(x) = \mathcal{F}[s(x)],$$

where the mapping $s : \mathbb{R}^n \mapsto \mathbb{R}^m$ describes the simulation as a function of the parameters (or controls) x.

Examples

- Nuclear fission
- Modeling subsurface flow in the Hanford area
- Rietveld refinement of powder diffraction data
- Pumping rates for the de-contamination of groundwater

Argonne NATIONAL LABORATORY

Parameter Estimation Problems in Nuclear Fission

- $x \in \mathbb{R}^n$ is the vector of parameters
- m is the number of nuclei.
- $f_k(x)$ is the vector of observables for the k-th nucleus.
- y_k is the data vector associated with the k-th nucleus.

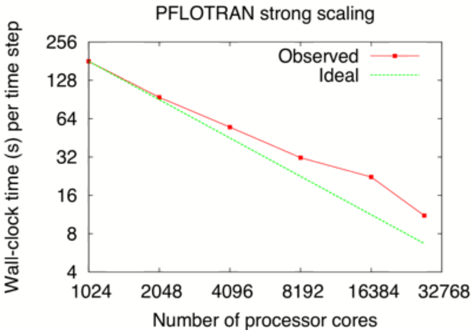
$$x \Longrightarrow \boxed{\text{HFBTHO, MFD, \ldots}} \Longrightarrow f_k(x)$$

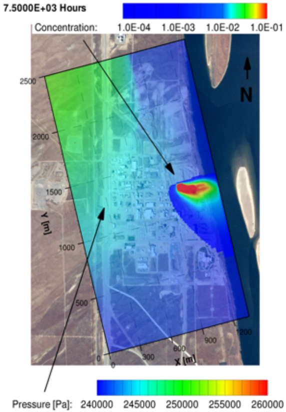
The least-squares approach (χ^2) requires the minimization of

$$f(x) = \frac{1}{2} \sum_{k=1}^{m} \sigma_k \| f_k(x) - y_k \|^2$$

where $\|\cdot\|$ is the l_2 norm and σ_k is a set of weights.

Subsurface Flow in the Hanford Area





- PFLOTRAN (1024 processors, Cray XT3) requires 2.3 hours
- Three layers and about two parameters per layer

Challenges of Complex Simulations

- Expensive evaluations of $f(x)$
- Large memory requirements
- Noisy function evaluations
- Lack of derivatives with respect to parameters
- Possibly several minima
- Limited computational budget

In a complex simulation, we need to obtain acceptable solutions in a small multiple of n function evalutions.

Derivative-Free Optimization Algorithms

Direct Search Methods

- Nelder-Mead (nmsmax, nelder, fminsearch)
- Pattern search (appspack, mdsmax, sid-psm, nomad)

Model-Based Methods

- Quadratic models (uobyqa, newuoa, dfo)
- Radial-basis models (orbit, boosters)
- Quasi-Newton models(imfil, . . .)

Numerical Recipes only mentions the Nelder-Mead method

Model-Based Derivative-Free Algorithms

For $k = 0, \ldots$

- ◇ Compute the model q_k.
- ◇ Check the fidelity of the model.
- ◇ Compute a step s_k in the trust-region $\|s\| \le \Delta_k$.
- ◇ Update x_k and the trust-region radius Δ_k.

Two possible classes of models

- ◇ Quadratic (full, partial, linear)
- ◇ Adaptive models

Quadratic Models

Quadratic models in $\mathbb{R}^n$

$$q(x) = c + g^T(x - x_0) + \tfrac{1}{2}(x - x_0)^T G(x - x_0)$$

are obtained by interpolating f at $\frac{1}{2}(n+1)(n+2)$ points.

Cost of computing a full quadratic model of f in $\mathbb{R}^n$

n	5	10	20	30	40	50
evaluations of f	21	66	231	496	861	1326

Minimal Norm (Frobenius) Quadratics

Least-change (partial) quadratic models determine a new quadratic q_+ for any $m \geq n+1$ points $x_1, \ldots x_m$ by requiring that

$$\min \left\{ \|\nabla^2 q_+ - \nabla^2 q\|_F^2 : q_+(x_k) = f(x_k), \quad 1 \leq k \leq m \right\}$$

where q is the current model.

- Existence of the model is not guaranteed
- There is at most one model that satisfies these conditions
- Computing the model requires order $(m+n)^3$ operations
- Model fidelity depends on the geometry of the sample points

Extensions

In least-squares problems where

$$f(x) = \frac{1}{2} \sum_{k=1}^{m} f_k(x)^2,$$

we can construct quadratic models $q_1, \ldots q_m$, and consider the (non-quadratic) model

$$f_M(x) = \frac{1}{2} \sum_{k=1}^{m} q_k(x)^2.$$

⚠ See Stefan Wild's talk for computational results.

Benchmarking

A **benchmark** is defined by

- ◇ $\mathcal{P}$ – Benchmark problems
- ◇ $\mathcal{T}$ – Convergence test
- ◇ $\mathcal{S}$ – Set of solvers

For each solver $s \in \mathcal{S}$, record the number of function evaluations required to *solve* each problem $p \in \mathcal{P}$.

- ◇ What is an appropriate set of benchmark problems?
- ◇ What is an appropriate convergence test?

J. Moré and S. Wild, *Benchmarking derivative-free optimization algorithms*, www.mcs.anl.gov/~more/dfo

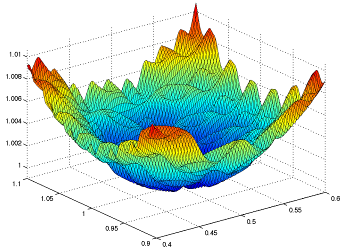
Argonne NATIONAL LABORATORY

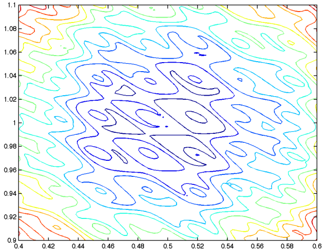
Example: Noisy Quadratics

Given a noise function $\phi : \mathbb{R}^n \mapsto [-1, 1]$, consider

$$f(x) = (1 + \tfrac{1}{2}\|x - e\|^2)(1 + \varepsilon_f \phi(x)),$$

where $\varepsilon_f = 10^{-3}$ is the noise level, and e is the vector of all ones.





Simulation-Based Optimization Problems

⚠ What kind of simulation do you have?

Problem	Simulation
Gaussian elimination	GE with row pivoting
Singular value estimation	condition number estimator
The non-negative svd	svd of a projection of a matrix
Particles with springs	eigenvalues of a symmetric matrix
Satellite orbit	ode23
Isomerization of α-pinene	ode45
Incompressible elastic rods	ode45

Joint work with J. Reed (UCSD)

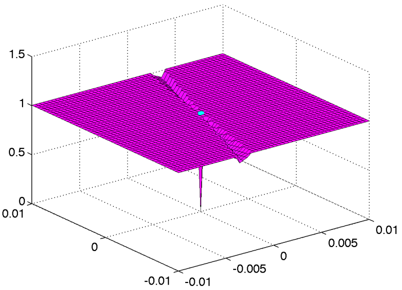
Simulation Slices: Singular Value Estimation

Find an upper triangular matrices $A \in \mathbb{R}^{d \times d}$ that minimize

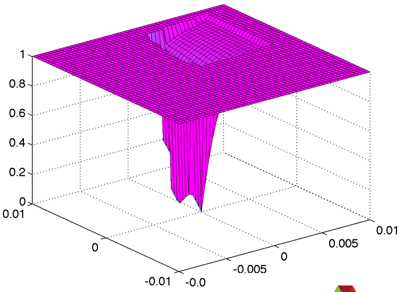
$$f(A) = \frac{\sigma_1(A)}{\text{estsv}(A)}$$

where $\text{estsv}(\cdot)$ is an estimate of σ_1.

Plot of f for the estsv problem with f = 1.17e-07

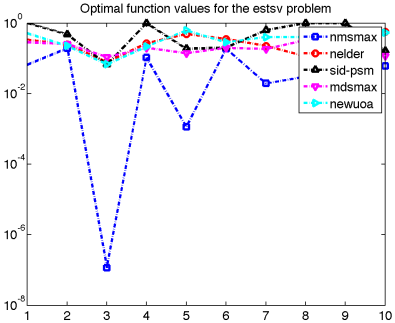


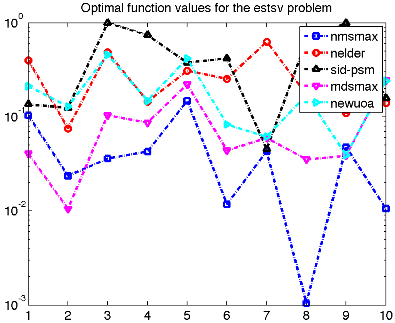
Plot of f for the estsv problem with f = 1.04e-03



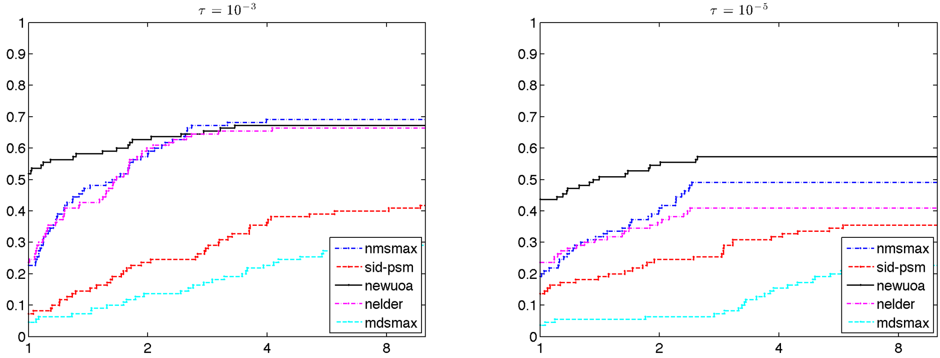
Argonne NATIONAL LABORATORY

Performance of Optimization Solvers: Singular Value Estimation





Performance Profiles for SBO Problems



Concluding Remarks

- Parameter estimation of complex simulations will require (extremely) advanced architectures
- Our benchmarking results show that model-based solvers generally perform best on complex simulations
- See Stefan Wild's talk for computational results on a nuclear fission (SciDAC UNEDF) application
- Talk to us for your optimization needs