Development of mathematical models for Exascale and beyond

I. S. Duff (RAL, UK and CERFACS, France)
S. Gratton (INPT-IRIT, University of Toulouse and CERFACS, France)
D. Titley-Peloquin (CERFACS, France)
X. Vasseur (CERFACS, France)

Examath 2013
Workshop on Applied Mathematics Research for Exascale Computing
August 21-22th, 2013, Washington D.C., USA
Introduction

Exascale stimulates the emergence of difficult problems to solve

Targeted long-term scientific challenges at CERFACS

- Multiphysics multi-scale coupled problems (fluid mechanics, combustion, turbomachinery)
- Stochastic, non-Gaussian, nonconvex optimization problems (data assimilation in meteorology or reservoir modeling)
- Huge-scale inverse problems in geosciences (seismic imaging)

New mathematical models at the interface between numerical analysis and statistics/probability have to be developed
Stochastic conditioning (I/II)

How sensitive are matrix functions to random noise in the data? Illustration with $f(A) = A^{-1}b$ with $A \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}^n$

- Normwise relative condition number of $f(A)$

$$\kappa_f(A) = \lim_{\varepsilon \to 0} \sup_{\|H\|_2 \leq \varepsilon} \frac{\|f(A + H) - f(A)\|_2/\|f(A)\|_2}{\|H\|_2/\|A\|_2}$$

- Stochastic alternatives to the usual normwise relative condition number of $A^{-1}b$ when $A$ is perturbed to $A + H$ with $\text{vec}(H) \sim \mathcal{N}(0, \sigma^2 I_{n^2})$

- For any non singular $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$, $\tau > 0$, independently of $\sigma$ [Sankar et al, 2006]:

$$\text{Prob}\left\{ \frac{\|(A + H)^{-1}b - A^{-1}b\|_2}{\|A^{-1}b\|_2} > \tau \right\} \leq \sqrt{\frac{8n}{\pi}} \frac{\|A^{-1}\|_2 \|b\|_2}{\tau \|A^{-1}b\|_2}$$

- This bound is independent of $\sigma$ but does depend on $n$

How sensitive are matrix functions to random noise in the data? Illustration with $f(A) = A^{-1}b$

- **Large asymptotics:**

  $$\text{Prob} \left\{ \left| \frac{\| (A + H)^{-1}b - A^{-1}b \|_2}{\| A^{-1}b \|_2} - 1 \right| > \tau \right\} \leq 2 \sqrt{\frac{2}{\pi}} \frac{\| A \|_2}{\tau \sigma}$$

- The relative error converges in probability to 1 as $\sigma \to +\infty$.

- **Illustration:** given samples of $\text{vec}(H) \sim \mathcal{N}(0, \sigma^2 I_{n^2})$ for several values of $\sigma$, we plot versus $\sigma$:
Numerical nonlinear optimization (I/II)

Combining evolutionary strategies (ES) and derivative free optimization (DFO) methods to address the solution of unconstrained optimization problems $\min_{x \in \mathbb{R}^n} f(x)$

- **Main goal**: to increase the exploration possibilities of the DFO algorithm thanks to populations based on random sampling

- **Global** convergence theory developed on a variant that both controls the length of the stepsize ($\sigma_k$) and imposes a sufficient decrease condition on the objective function

\[
f(x_{k+1}) \leq f(x_k) - \rho(\sigma_k)
\]

- Problem with up to several hundreds of variables has been successfully solved and an extension has been provided to tackle constrained optimization problems.

Numerical nonlinear optimization (II/II)

Data profiles on Moré-Wilde testcases (constrained optimization problems)

- Population based method based on CMA-ES [Hansen, 1996]

- Improvements in efficiency without deteriorating the behavior of CMA-ES in the presence of nonconvexity

- Excellent results compared with PSwarm [Vaz et al, 2007].
Numerical linear algebra (I/III)

Subset of core problems to address for Exascale

▶ Analysis of the convergence of inner-outer algorithms that minimize data movement (variable multilevel preconditioner, flexible Krylov subspace methods, inexact Krylov subspace methods, embedded iterations in hybrid direct-iterative methods or eigenvalue problems, stopping criteria),

▶ Design data compression algorithms in sparse direct methods with well controlled accuracy (e.g. low-rank approximations),

▶ Develop error analysis in standard algorithms when $n \to +\infty$ or when multi-precision is used,

▶ Application framework: full waveform inversion in seismic: solution of heterogeneous Helmholtz problems

Multilevel preconditioners for Helmholtz problems (II/III)

Salt dome (40 Hz), $3711 \times 3711 \times 1149$ (16384 cores of BG/P)

- Efficiency of a new preconditioner applied to the Helmholtz operator on the finest level and based on the complex shifted Laplacian idea [Erlangga et al., 2006] to obtain approximate coarse level solutions
Error analysis in Gaussian elimination (III/III)

Behaviour of the error bound in the Gaussian elimination when $n \to +\infty$

- Backward error bounds typically involve a moderate function of the problem size $n$ ($A \in \mathbb{R}^{n \times n}$):

  $$(A + \Delta A) = PLU \quad \|\Delta A\|_\infty \leq 8n^3 \rho_n(A) \|A\|_\infty \varepsilon + O(\varepsilon^2)$$

- Need error bounds or estimates that remain meaningful as $n \to +\infty$. 
Bibliography I

S. Gratton and D. Titley-Peloquin
*Sensitivity of systems of equations to random noise.*

S. Gratton and D. Titley-Peloquin
*Stochastic conditioning of matrix functions.*

Y. Diouane, S. Gratton and L. N. Vicente
*Globally convergent evolution strategies and CMA-ES.*

H. Calandra, S. Gratton, X. Pinel and X. Vasseur
*An improved two-grid preconditioner for the solution of three-dimensional Helmholtz problems in heterogeneous media.*

Thank you for your attention!