Workshop on Quasi-Newton Methods

Book of Abstracts

9th November 2022, Edinburgh

THE UNIVERSITY of EDINBURGH

Organizing Committee

- **Cipolla Stefano** (The University of Edinburgh, School of Mathematics)
- **Gondzio Jacek** (The University of Edinburgh, School of Mathematics)

Contents

Workshop on Quasi-Newton Methods

Schedule

The program below refers to the UK time zone (UTC+0)

Nov 9th, 2022	
13.00 ⟶ 13.05 **Opening Remarks**	
13.00 ⟶ 15.00 Chair: S. Cipolla	
13.05 ⟶ 14.00	L. Bergamaschi (ERGO Speaker)
14.00 ⟶ 14.45	F. Sobral
14.45 ⟶ 15.10 Coffee Break	
15.10 ⟶ 17.30 Chair: J. Gondzio	
15.10 ⟶ 16.05	J. Nocedal (Keynote Speaker)
16.05 ⟶ 16.50	F. Mannel
16.50 ⟶ 17.35	S. Cipolla
18.30 Social Dinner: The Rabbit Hole	

Abstracts

Quasi-Newton preconditioners, motivation and theory

Luca Bergamaschi[1]

[1] Department of Civil Environmental and Architectural Engineering, University of Padova, Italy. luca.bergamaschi@unipd.it

Nonlinear problems and optimization

Solution to large and sparse non linear systems of equations, $F(x) = 0$, by Newton's method, require the repeated solution of linear systems with the same structure. If iterative solvers are selected, they need to be accelerated by suitable preconditioners. Also optimization problems, written as $\min f(x)$, fall in this framework as they are commonly tackled by finding the zero of the gradient of the objective function.

In this talk we assume that the Jacobian J of F (or the Hessian H of f) is known, either explicitly, or as the results of its application of a vector. Our concern is to accelerate the Krylov subspace iterative solution to the linear system at each step of the Newton sequence [8].

To solve the Newton systems

$$J(x_k) = -F(x_k), k = 0, \ldots,$$

we define sequences of preconditioners

$$B_0 \approx J(x_0)^{-1}, \ldots, B_k, \ldots$$

using the well-known Quasi-Newton updates, such as Broyden and BFGS, in case of non-symmetric and symmetric positive definite Jacobians, respectively.

We review the theoretical properties of these sequence of preconditioners (starting from the well known *bounded deterioration property*) which aim at maintaining controlled the condition number of the preconditioned matrices in the sequence [4, 2]. In view of the sparsity of the matrices involved, we develop a (limited memory variant of a) vector recurrence formula for the application of the preconditioner which (apart of $o(n)$ operations), involve the solution of a number of linear systems with the same (B_0) matrix, which can factored (or preconditioned) once and for all.

Results on various realistic nonlinear and optimization problems are presented [3, 5] which show the acceleration of convergence produced by employing such a sequence of preconditioners.

Sequences of linear systems

The Quasi-Newton update formulae are all based on two Newton vectors, $s_k = x_{k+1} - x_k$ and $y_k = F(x_{k+1}) - F(x_k)$ used to mimic the secant condition. In case however of a sequence of linear systems $Ax = b_k, k = 1, \ldots, N$, not necessarily arising from a linearization, the previously considered updated can be used by making the substitutions $s_k \longrightarrow v$ (with v arbitrary) and $y_k \longrightarrow Av$. This change gives raise to the *low rank update of preconditioners* which may speed up enormously (and save on the cost of computing a preconditioner at each system in the sequence) the iterative solution of the k-th system [1, 6]. Remarkably these update formulae have been successfully used in the past without connecting them with the Quasi-Newton ideas as e.g. the *balancing preconditioner* which exploits the BFGS idea in [9] or the *tuned preconditioner* proposed in [7], which is nothing but the SR1 update formula, in the framework of iterative eigensolvers.

References

[1] L. Bergamaschi, *A survey of low-rank updates of preconditioners for sequences of symmetric linear systems*, Algorithms, 34 (2) (2020).

[2] L. Bergamaschi, R. Bru, and A. Martínez, *Low-rank update of preconditioners for the inexact Newton method with SPD jacobian*, Mathematical and Computer Modelling, 54 (2011), pp. 1863–1873.

[3] L. Bergamaschi, R. Bru, A. Martínez, J. Mas, and M. Putti, *Low-rank update of preconditioners for the nonlinear Richard's equation*, Mathematical and Computer Modelling, 57 (2013), pp. 1933–1941.

[4] L. Bergamaschi, R. Bru, A. Martínez, and M. Putti, *Quasi-Newton preconditioners for the inexact Newton method*, Electron. Trans. Numer. Anal., 23 (2006), pp. 76–87.

[5] L. Bergamaschi, R. Bru, A. Martínez, and M. Putti, *Quasi-Newton acceleration of ILU preconditioners for nonlinear two-phase flow equations in porous media*, Advances in Engineering Software, 46 (2012), pp. 63–68.

[6] L. Bergamaschi, V. De Simone, D. di Serafino, and A. Martínez, *BFGS-like updates of constraint preconditioners for sequences of KKT linear systems*, Numer. Linear Algebra Appl., 25 (2018), pp. 1–19. e2144.

[7] M. A. Freitag and A. Spence, *A tuned preconditioner for inexact inverse iteration applied to Hermitian eigenvalue problems*, IMA J. Numer. Anal., 28 (2008), pp. 522–551.

[8] J. L. Morales and J. Nocedal, *Automatic preconditioning by limited memory Quasi-Newton updating*, SIAM J. Optim., 10 (2000), pp. 1079–1096.

[9] R. Nabben and C. Vuik, *A comparison of deflation and the balancing preconditioner*, SIAM J. Sci. Comput., 27 (2006), pp. 1742–1759.

Quasi-Newton: from *low complexity matrix algebras* to Anderson Acceleration.

Stefano Cipolla[1]

[1] School of Mathematics, The University of Edinburgh, UK. scipolla@exseed.ed.ac.uk

Quasi-Newton methods continue to have an undeniable charm for computational scientists and play, indeed, a central role in various branches of optimization and numerical analysis. Aim of this talk is to present a series of recent results on Quasi-Newton methods in connection with apparently different areas of computational science.

In particular, in the first part of this talk, we will present a series of results concerning the possibility of using *Householder adaptive transforms* to produce convergent Broyden Class-type methods having linear complexity [2].

In the second part, instead, we will show how Anderson Acceleration can be interpreted as a particular instance of a Quasi-Newton method and how this interpretation can be used to obtain local convergence results [1].

References

[1] C. Brezinski, S. Cipolla, M. Redivo-Zaglia, Y. Saad. Shanks' and Anderson-type techniques for systems of nonlinear equations. IMA J. Numer. Anal., in press, 2021. `https://doi.org/10.1093/imanum/drab061`.

[2] S. Cipolla, C. Di Fiore, and P. Zellini. A variation of Broyden Class methods using Householder adaptive transforms. Comput. Optim. Appl., 77(2), 433–463, 2020. doi.org/10.1007/s10589-020-00209-8.

Semismooth Quasi-Newton methods and other recent developments in Quasi-Newton methods

Florian Mannel[1]

[1] Institute of Mathematics and Image Computing University of Lübeck, Germany. `mannel@mic.uni-luebeck.de`

Quasi-Newton methods continue to be an active area of research. For instance, it is an ongoing effort to understand how to effectively use Quasi-Newton methods for solving nonsmooth problems. In the main part of this talk I show that Broyden's method can be applied to the smooth parts of an overall semismooth problem. In fact, by combining Broyden's method with a semismooth Newton method we obtain an algorithm that dramatically outperforms state-of-the-art semismooth Newton methods in a large-scale real-world application from optimal control. We prove local superlinear convergence in an infinite dimensional setting, which allows to deduce convergence results in various function spaces. Notably, these results accurately reflect the convergence behaviour that is observed in practice.
I will also present a globalized version of this algorithm for generalized complementarity problems.

In the second part of the talk we briefly discuss some other current research topics in Quasi-Newton methods. Here, I plan to address

- the order of convergence of Broyden's method,
- Broyden's method for singular problems,
- Quasi-Newton methods in machine learning.

Throughout the talk I point out some open problems as avenues for future research and potential collaboration.

References

[1] F. Mannel, A. Rund. *A hybrid semismooth Quasi-Newton method for nonsmooth optimal control with PDEs*, Optim. Eng. 22 (2021), 2087–2125.

[2] F. Mannel, A. Rund. *A hybrid semismooth Quasi-Newton method for structured nonsmooth operator equations in Banach spaces*, J. Convex Anal. 29 (2022), 183–204.

Quasi-Newton Methods in the Presence of Noise

Jorge Nocedal[1]

[1] Robert R. McCormick School of Engineering and Applied Science, Northwestern University, USA. nocedal@ece.northwestern.edu

Classical analysis of quasi-Newton methods assumes that function and gradient evaluations are exact. However, when these evaluations are noisy the standard BFGS method may fail. We describe a simple modification of BFGS with an Armijo-Wolfe line search and show that it is guaranteed to converge to a neighborhood of the solution that is determined by the size of the noise. Our results extend the existing analysis on the behavior of BFGS updating to the stochastic setting. We present numerical results illustrating the performance of the new BFGS method in the presence of bounded noise.

Polynomial worst-case complexity for Quasi-Newton Primal Dual Interior Point methods

Francisco Sobral[1]

[1] Department of Mathematics State University of Maringá, Brazil. fncsobral@uem.br

In this talk we present polynomial worst-case complexity results for a simplified class of Quasi-Newton Primal Dual Interior Point Methods in linear programming problems. The feasible and infeasible cases are considered and well-known neighborhoods of the central path are used. To the best of our knowledge, those are the first complexity results for this class of algorithms in literature.

Author Index