MONDAY 11 April

Dr Serge Gratton (INP-ENSEEIHT, Toulouse)

Data Assimilation Recurrent networks can beat Ensemble methods

TBC

Professor Patrick Farrell (University of Oxford)

Reynolds-robust preconditioners for the stationary incompressible Navier-Stokes equations

When approximating PDEs with the finite element method, large sparse linear systems must be solved. The ideal preconditioner yields convergence that is algorithmically optimal and parameter robust, i.e. the number of Krylov iterations required to solve the linear system to a given accuracy does not grow substantially as the mesh or problem parameters are changed.

Achieving this for the stationary Navier-Stokes has proven challenging: LU factorisation is Reynolds-robust but scales poorly with degree of freedom count, while Schur complement approximations such as PCD and LSC degrade as the Reynolds number is increased.

Building on the work of Schöberl, Olshanskii, and Benzi, in this talk we present the first preconditioner for the Newton linearisation of the stationary Navier--Stokes equations in three dimensions that achieves both optimal complexity and Reynolds-robustness. The exact details of the preconditioner varies with discretisation, but the general theme is to combine augmented Lagrangian stabilisation, a custom multigrid prolongation operator involving local solves on coarse cells, and an additive patchwise relaxation on each level that captures the kernel of the divergence operator.

We present 3D simulations with over one billion degrees of freedom with robust performance from Reynolds number 10 to 5000. We also present recent extensions to implicitly-constituted non-Newtonian problems, and to magnetohydrodynamics.

Professor Jennifer Scott (The University of Reading and STFC Rutherford Appleton Laboratory)

Can randomised preconditioning improve the weather forecast?

We consider the large sparse weighted least squares problems that arise in the solution of weak constraint four-dimensional variational data assimilation, a method of significant interest for numerical weather prediction. In this talk, we focus on preconditioning the normal equations. This is challenging because of the size of the system, the cost of products with the system matrix, and the need to severely limit the number of iterations of the CG solver to ensure a forecast is obtained in a timely manner. Exploiting the recent resurgence of randomised methods, we propose using randomised methods to develop limited memory preconditioners that are simple to compute and to apply. We illustrate their effectiveness using a model problem and look at how the number of observations of the dynamical system influences performance.

This is joint work with Ieva Dauzickaite, Amos Lawless, and Peter Jan van Leeuwen.

Ms Ieva Dauzickaite (Charles University)

Randomised preconditioning for saddle point systems in data assimilation

Saddle point formulations of linear systems of equations occurring in the incremental weak constraint four-dimensional variational data assimilation method are suitable for time-parallel computations. These large sparse systems are solved using Krylov subspace solvers and preconditioning is needed to do it efficiently. We consider block diagonal preconditioner that employs an approximation of the Schur complement. Observation information is usually excluded in the approximation and we propose a new way to incorporate it. This can be achieved by computing a randomised eigenvalue decomposition. An idealised numerical example shows that the preconditioner is effective.

This is joint work with Jennifer Scott, Amos Lawless, and Peter Jan van Leeuwen.

Dr Aretha Teckentrup

Convergence and Robustness of Gaussian Process Regression

We are interested in the task of estimating an unknown function from data, given as a set of point evaluations. In this context, Gaussian process regression is often used as a Bayesian inference procedure, and we are interested in the convergence as the number of data points goes to infinity. Hyper-parameters appearing in the mean and covariance structure of the Gaussian process prior, such as smoothness of the function and typical length scales, are often unknown and learnt from the data, along with the posterior mean and covariance. We work in the framework of empirical Bayes', where a point estimate of the hyper-parameters is computed, using the data, and then used within the standard Gaussian process prior to posterior update. Using results from scattered data approximation, we provide a convergence analysis of the method applied to a fixed, unknown function of interest.

[1] A.L. Teckentrup. Convergence of Gaussian process regression with estimated hyper-parameters and applications in Bayesian inverse problems. SIAM/ASA Journal on Uncertainty Quantification, 8(4), p. 1310-1337, 2020.

Dr. Stefano Cipolla (The University of Edinburgh)

Primal Dual Regularized IPM: a Proximal Point perspective

Computational evidence suggests that the Primal-Dual Regularization for Interior Point Methods (IPMs) is a successful technique able to stabilize and to speed-up the linear algebra used in IPM implementations [1]. On the other hand, many issues remain open when IPMs are used in their primal-dual regularized form and, in particular, to the best of our knowledge, the known convergence theory requires strong assumptions on the uniform boundedness of the Newton directions [2]. Recently, the study of the interaction of primal-dual regularized IPMs with the Augmented Lagrangian Method and the Proximal Point Algorithm has permitted to prove the convergence when the regularization parameter is driven to zero at a suitable speed [3].

In this talk, we will show that it is possible to naturally frame the primal-dual regularized IPMs in the context of the Proximal Point Algorithm [4]. Among the benefits of the proposed approach, we will show how convergence can be guaranteed without any supplementary assumptions and how the rate of convergence can be explicitly estimated in relation to (fixed) regularization parameter. Additionally, numerical results proving the efficiency and reliability of the proposed approach will be presented.

References

[1] Altman, A., Gondzio, J. (1999) Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization, Optimization Methods and Software, 11:1-4, 275-302.

[2] Friedlander, M. P., Orban D. (2012) A primal-dual regularized interior-point method for convex quadratic programs, Mathematical Programming Computation 4.1 : 71-107.

[3] Pougkakiotis, S., Gondzio J. (2021) An interior point-proximal method of multipliers for convex quadratic programming, Computational Optimization and Applications 78.2 : 307-351.

[4] Cipolla, S., Gondzio J. (In preparation) Proximal stabilized Interior Point Methods for quadratic programming and low-frequency-updates preconditioning techniques.

TUESDAY 12 April

Dr Joanne Waller (Met Office)

Can we handle spatial observation error correlations in atmospheric data assimilation?

Over the last 10 years the use of inter-channel correlations in data assimilation has resulted in improved weather forecasts. However, despite the existence of significant spatial error correlations they are not yet widely accounted for in atmospheric data assimilation.

This presentation will consider how we can handle these spatial error correlations. First, the importance and expected benefits of including spatial error correlations will be discussed. Next, potential methodology for handling spatial correlations will be presented along with some of the technical challenges that still remain.

Finally, we present results that show the impact of assimilating Doppler radar radial wind observations when explicitly accounting for the correlated observation errors. We conclude that, though some technical challenges remain, we are able to handle some spatial observation error correlations, and the benefits of this are beginning to emerge.

Dr. Selime Gurol (CERFACS)

Latent Space Data Assimilation

Computationally efficient Data Assimilation (DA) methods are crucial in Earth system modeling, particularly at the time of big data where huge quantities of observations are available.

Capitalizing on the ability of Neural Networks techniques for approximating the solution of PDE's, we incorporate Deep Learning (DL) methods into a DA framework. More precisely, we exploit the latent structure provided by autoencoders (AEs) to design a DA methodology in the latent space.

Model dynamics are also propagated within the latent space via a surrogate neural network. This latent space DA algorithm is tested on an augmented version of Lorenz 96 equations such that the augmented system possesses a latent structure that accurately represents the observed dynamics.

Numerical experiments based on this particular system evidence that exploiting the latent structure of the problem can reduce the computational cost and provide better accuracy than state-of-the-art algorithms.

Dr. Hussam Al Daas (Rutherford Appleton Laboratory, STFC)

Recent advances in robust algebraic domain decomposition preconditioners

Solving sparse linear systems is omnipresent in scientific computing. Direct approaches based on matrix factorization are very robust, and since they can be used as a black-box, it is easy for other software to use them. However, the memory requirement of direct approaches scales poorly with the problem size, and the algorithms underpinning sparse direct solvers software are poorly suited to parallel computation.

Multilevel Domain decomposition (MDD) methods are among the most efficient iterative methods for solving sparse linear systems.

One of the main technical difficulties in using efficient MDD methods (and most other efficient preconditioners) is that they require information from the underlying problem which prohibits them from being used as a black-box. This was the motivation to develop the widely used algebraic multigrid for example.

I will present a series of robust and fully algebraic MDD methods, i.e., that can be constructed given only the coefficient matrix and guarantee a priori prescribed convergence rate.

The series consists of preconditioners for least-squares problems, sparse SPD matrices, general sparse matrices (two), and saddle-point systems.

Numerical experiments illustrate the effectiveness, wide applicability, scalability of the proposed preconditioners.

A comparison of each one against state-of-the-art preconditioners is also presented.

Dr Jonas Latz (Heriot-Watt University, Edinburgh)

Fast and even faster sampling of parameterised Gaussian random fields

Gaussian random fields are popular models for spatially varying uncertainties, arising, e.g., in geotechnical engineering, hydrology, or image processing. A Gaussian random field is fully characterised by its mean and covariance operator. In more complex models these can also be partially unknown. In this case we need to handle a family of Gaussian random fields indexed with hyperparameters. Sampling for a fixed configuration of hyperparameters is already very expensive due to the nonlocal nature of many classical covariance operators. Sampling from multiple configurations increases the total computational cost severely. In this talk we employ Karhunen-Loève expansions and adaptive cross approximations for sampling.

To reduce the cost we construct a reduced basis surrogate built from snapshots of Karhunen-Loève eigenvectors in the first case. In the second case, we propose a parameterised version of the adaptive cross scheme. In numerical experiments we consider Matérn-type covariance operators with unknown correlation length and standard deviation. Here, we study the approximation accuracy of reduced basis and cross approximation. As an application we consider Bayesian inversion with an elliptic partial differential equation where the logarithm of the diffusion coefficient is a parameterised Gaussian random field. Indeed, we employ Markov chain Monte Carlo on the reduced space to generate samples from the posterior measure.

Mr Filippo Zanetti (University of Edinburgh)

New indicators for the early termination of the linear solver in interior point methods

When an iterative method is applied to solve the linear equation system in Interior Point Methods (IPMs), the attention is usually placed on accelerating their convergence by designing appropriate preconditioners, but the linear solver is applied as a black box solver with a standard termination criterion which asks for a sufficient reduction of the residual in the linear system. Such an approach often leads to an unnecessary ''oversolving'' of linear equations. In this talk, it is shown how an IPM can preserve the polynomial worst-case complexity when relying on an inner termination criterion that is not based on the residual of the linear system. Moreover, a practical criterion is derived from a deep understanding of IPM needs. The new technique has been adapted to the Conjugate Gradient (CG) and to the Minimum Residual method (MINRES) applied in the IPM context. The new criterion has been tested on a set of quadratic optimization problems including compressed sensing, image processing and instances with partial differential equation constraints, and it has been compared to standard residual tests with variable tolerance. Evidence gathered from these computational experiments shows that the new technique delivers significant improvements in terms of inner (linear) iterations and those translate into significant savings of the IPM solution time.

Miss Lingyi Yang (University of Oxford)

Path signatures and neural controlled differential equations for prediction tasks

When working with real data, despite the fact that some variables are inherently continuous, we only observe them at discrete intervals. As a result, we need to find a way to interpret new data releases as coming from some underlying continuous time series and exploit connections of the inputs to best model the target variable. The path signature is a natural feature set to consider for time-series data, and more generally high dimensional paths. In practice, our data may be irregular, and have mixed frequency or be missing across the different variables measured. By interpreting the time series as a path object, all of these problems can be handled.

Differential equation-inspired neural networks have become increasingly popular. Motivated by controlled differential equations, neural controlled differential equations (Neural CDEs) can be thought of as a continuous-time extension of recurrent neural networks (RNNs). This requires the construction of a continuous control path from discrete observations. For online prediction tasks, the underlying control path should be causal, something that was not considered when Neural CDEs were first conceived.

In this talk, we highlight practical uses of signatures in prediction tasks. We discuss how to obtain a causal control path, along with other properties that the ideal control path should satisfy. Finally we show that this modification for neural CDEs achieves state-of-the-art performance on prediction problems from the MIMIC-IV medical dataset.

Jonna Roden (Online Lecture), University of Edinburgh

Spectral element methods and iterative solvers for PDE-constrained optimization problems

In this talk I will introduce a numerical method to solve optimal control problems with (integro-)PDE constraints. This method combines spectral elements with a Newton-Krylov algorithm, which provides a tool for the fast and accurate solution of the resulting optimality systems. In particular, this framework allows for the solution of (integro-)PDE models and optimal control problems on complex domains, which is a crucial feature in accurately describing various (industry) applications. Finally, some examples of current work and future industrial applications will be given. This is joint work with Ben Goddard and John Pearson.

WEDNESDAY 13 April

Dr Alison Ramage (University of Strathclyde)

Preconditioning for Data Assimilation Problems

Traditionally, large-scale variational data assimilation problems have been commonly found in applications like numerical weather prediction and oceanographic modelling, where the 4D-Var method is frequently used to calculate a forecast model trajectory that best fits the available observations to within the observational error over a period of time.
However, due to memory limitations, in practice it is often impossible to assemble, store or manipulate the underlying matrices explicitly for a realistic model, as the state vectors used in such applications could contain billions or trillions of unknowns.

Using the 4D-Var method leads to an unconstrained optimisation problem, for which a limited-memory representation of the Hessian can be built based on a small (compared to the dimension of the state vector) number of Hessian eigenvalues, computed by the Lanczos method. In this work, we discuss an building a limited-memory approximation of the inverse Hessian of the linearised quadratic minimisation subproblems, again using the Lanczos method. We apply this idea as a preconditioner within 4D-Var via a multilevel structure, and show that this can reduce overall memory requirements and increase computational efficiency.

Dr Silvia Gazzola (University of Bath)

Regularization by inexact Krylov methods

This talk will present theoretical and algorithmic aspects of regularization methods based on inexact Krylov methods for the solution of large-scale discrete inverse problems. Specifically, we will introduce two new inexact Krylov methods that can be efficiently applied to unregularized or Tikhonov-regularized least squares problems, and we present their theoretical properties, including links with their exact counterparts and strategies to monitor the amount of inexactness.

We then describe how the new methods can be applied to solve separable nonlinear inverse problems arising in blind deblurring, where both the sharp image and the parameters defining the blur are unknown. We show that the new inexact solvers (which can naturally handle varying inexact blurring parameters while solving the linear deblurring subproblems within a variable projection method) allow for a much reduced number of total iterations and substantial computational savings with respect to their exact counterparts.

Dr. Davide Palitta (Alma Mater Studiorum, Universita' di Bologna)

Stein-based Preconditioners for Weak-constraint 4D-var

Algorithms for data assimilation try to predict the most likely state of a dynamical system by combining information from observations and prior models. One of the most successful data assimilation frameworks is the linearized weak-constraint four-dimensional variational assimilation problem (4D-Var), that can be ultimately seen as a minimization problem. One of the main challenges of such approach is the solution of large saddle point linear systems arising as inner linear step within the adopted nonlinear solver. The linear algebraic problem can be solved by means of a Krylov method, like MINRES or GMRES, that needs to be preconditioned to ensure fast convergence in terms of number of iterations. In this talk we will illustrate novel, efficient preconditioning operators which involve the solution of certain Stein matrix equations. In addition to achieving better computational performance, the latter machinery allows us to derive tighter bounds for the eigenvalue distribution of the preconditioned saddle point linear system.

A panel of diverse numerical examples displays the effectiveness of the proposed methodology compared to current state-of-the-art approaches.

The results presented in this talk come from a joint work with J. Tabeart.