"Linearisation techniques for solving mixed 0-1 quadratic programming problems" ..... Dr J. Lazić, MathWorks[10]

Mixed Integer Quadratic Programming problem (MIQP) involves minimising a quadratic objective function, subject to linear constraints, where some of the variables have to be integer (see, for example [5, 7,9]). A special case of MIQP is Mixed 0-1 Quadratic Programming problem where, in addition, all integer variables can only take values 0 or 1 (see [2,7]).

Mixed 0-1 Quadratic Programming problem has numerous applications in a variety of areas, ranging from finance to computational biology (see, for example [3,4,6]). However, such problems are known to be NP-hard (see [7,8]) and many examples still remain hard to solve, even with the rapid development of computing resources. A common way of tackling Mixed 0-1 Quadratic Programming problem is to reformulate it to an equivalent Mixed Integer Linear Programming (MILP) problem and then apply MILP solution techniques (see [1]).

This project is sponsored by MathWorks the purveyor and developer of MATLAB software. Starting from a review of existing linearisation methods for solving Mixed 0-1 Quadratic Programming problem , the goal of the project will be to compare several linearisation methods implemented in MATLAB and report the results, summarising the performance of the methods, with respect to solution quality (as compared to optimal solution - if one exists) and computational time required. The student may decide to implement one or more linearisation techniques.

Subject to time permitting, the student will be welcome to suggest further improvements to some of those methods, beyond the current state-of-the-art.

In summary, the project can be divided into the following stages:

(a) literature review

(b) gathering problem instances

(c) implementation of solution method(s) in MATLAB

(d) performance comparison of different methods and summary of results

(e) suggestions for further research to improve the current state-of-the-art

References

[1] Adams, W.P., Forrester, R.J. & Glover, F.W., 2005. Comparisons and enhancement strategies for linearising mixed 0-1 quadratic programs. *Discrete Optimization*, 1 (**2**), 99–120.

[2] Bienstock, D., 1997. Computational study of a family of mixed-integer quadratic programming problems. *Mathematical programming*, 85.2, 121–150.

[3] Cesarone, F., Scozzari, A. & Tardella, F., 2014. A new method for mean-variance portfolio optimization with cardinality constraints. *Annals of Operations Research*, 206 (**1**), 214–245.

[4] Forrester, R.J. & Greenberg, H.J., 2009. Quadratic binary programming models in computational biology. *Algorithmic Operations Research*, 4 (**2**).

[5] Lazimy, R., 1992. Mixed-integer quadratic programming. *Mathematical Programming*, 22 (**1**), 442–459.

[10]Professor A. Iserles is the departmental contact for this project

Moment methods for population balance equations

Supervisor: Professor Markus Kraft, Department Chemical Engineering & Biotechnology
Maths contact: Professor James Norris

Consider the time evolution of a particle system described by Smoluchowski's equation [4]

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = \frac{1}{2}\sum_{j=1}^{i-1}\beta_{j,i-j}N_jN_{i-j} - \sum_{j=1}^{\infty}\beta_{i,j}N_iN_j, \qquad i,j \in \mathbb{N} \tag{1}$$

where N_i is the number density of particles of size i and $\beta_{i,j}$ is a kernel describing the rate of successful collisions between particles of size i and j.

The method of moments [2] is often used as a computationally efficient alternative to the direct solution of equation (1). However, the moment evolution equations derived from equation (1) are generally unclosed. Although several well-established closures exist, including quadrature methods [3] and interpolation between known moments [1], moment methods have never been successfully applied to describe the depletion of a population due to processes which remove mass from the particles. Such situations are of critical importance to many practical problems such as the oxidation of soot particles formed in modern Diesel engines.

The aim of this project is to extend the method of moments with interpolative closure [1] to include processes which describe the depletion of a population due to mass removal from the particles.

References

[1] M. Frenklach. Method of moments with interpolative closure. *Chemical Engineering Science*, 57(12):2229–2239, 2002. doi: 10.1016/S0009-2509(02)00113-6.

[2] H. M. Hulbert and S. Katz. Some problems in particle technology: A statistical mechanical formulation. *Chemical Engineering Science*, 19(8):555–574, 1964. doi: 10.1016/0009-2509(64)85047-8.

[3] R. McGraw. Description of aerosol dynamics by the quadrature method of moments. *Aerosol Science and Technology*, 27(2):255–349, 1997. doi: 10.1080/02786829708965471.

[4] M. v. Smoluchowski. Versuch einer mathematischen Theorie der Koagulationskinetik kolloidaler Lösungen. *Zeitschrift für Physikalische Chemie*, 92: 129–168, 1917.

Optimization techniques for neural nets [numerical]

Dr D K Waymont, Waymont Consulting Ltd

Maths contact: Dr Mark Spivack

In the last decade artificial neural nets have become practically useful, being used commercially in handwriting recognition (automatic letter sorting) and speech recognition. Applications are now being found in many other classification or recognition tasks (e.g. image labelling and fraud detection). In the near future it is likely that computer games and robotics will include more and larger neural nets: the problem then being the number of parameters.

The resurgence of neural nets is due to a number of reasons: the internet is a large and readily available source of training and test data; algorithms for many of the difficult tasks such as learning algorithms and how to implement short term memory have been found; and the computing power available is sufficient to train large nets (millions of weights) in a reasonable time (i.e. days). If the nets could be trained more rapidly or more accurately then progress in the field could be accelerated – it is this problem that will be explored in this project.

A neural net consists of nodes that take signals from many other nodes or the input, sum them according to a set of weights and apply a monotonic, usually non-linear function to the sum to produce an output. Training a neural net means finding the weights. During training, the information available is in the form of weight derivatives. There are standard batch optimization techniques that can be used if the full training set is used (e.g. conjugate gradients, LBFGS – see [1]) but if the data sets are large and highly redundant – which is often the case – it is quicker to train on a subset of the training data, a "mini-batch", and then use a different training set on the next mini-batch. This leads to a stochastic gradient which breaks the assumptions of the batch optimization techniques. A number of stochastic gradient techniques have been used [2,3]: back propagation (gradient descent), back propagation with (Nesterov) momentum, etc; and also a range of techniques for overcoming problems with widely differing component magnitudes: adaptive step sizes and scaling strategies (e.g. rprop, rmsprop).

The project will compare a new algorithm which only uses the gradient direction with existing algorithms to determine its strengths and weaknesses. The model problem will be to identifying hand-written digits in the MNIST database using single and multi-layer classifiers. The new method is based on a particular discretization of the following updates:

$$\frac{d\boldsymbol{s}}{dt} = \alpha|\boldsymbol{s}|\frac{\boldsymbol{g}}{|\boldsymbol{g}|}$$
$$\frac{d\boldsymbol{p}}{dt} = |\boldsymbol{s}|\left(\frac{\boldsymbol{s}}{|\boldsymbol{s}|} + \beta\frac{\boldsymbol{g}}{|\boldsymbol{g}|}\right)$$

where $\boldsymbol{g}$ is the gradient, $\boldsymbol{s}$ an internal search direction, and $\boldsymbol{p}$ the current position. The precise discretization will be supplied.

The project would suit a candidate with some programming experience.

[1] Press WH, Teukolsky SA, Vetterling WT, Flannery BP, "Numerical recipes", CUP

[2] Sutskever I, Martens J, Dahl G, Hinton G, "On the importance of initialization and momentum in deep learning", Proceedings of the 30th Conference on Machine Learning, Atlanta, USA, 2013.

[3] Schaul T, Zhiang S, LeCun Y, "No More Pesky Learning Rates", arXiv:1206:1106

Mini-project: Using dolphin whistles to identify species.
Dr P. Dobbins
maths contact: Dr Mark Spivack

Dolphin whistles form time-frequency contours that are complex, overlapping, and unpredictable [1]. These features make automatic detection, classification and localisation of whistles difficult and labour-intensive. However, all such whistles fall within a fixed range of time and frequency parameters [2]. It has previously been demonstrated that a dolphin whistle detector can be constructed in MATLAB using this information [3, 4] and it is hoped to build on the achievements of that project by identifying which of the thirty-three species of marine dolphins produced the whistle. The project will start with a brief literature search for dolphin whistle classification techniques and a review of the theory behind any that are found. Then it will proceed to collect and analyse a set of the previously recorded dolphin whistles available at various internet sites, to extract the time and frequency statistics and then to determine the differences between a limited selection of species. Given this information, the main challenge in this mini-project will be to develop a pattern recognition algorithm which identifies signals that match the specific set of parameters for a particular species, whilst rejecting interference such as ambient noise and transient sounds. This algorithm could be based on the previous MATLAB whistle detector [3], algorithms published in the literature, or could take an entirely new approach. Finally, the probability of detection and false alarm rate of the algorithm will be assessed using a different set of previously recorded whistles.

References

[1] Caldwell, M.C. and D.K. Caldwell, 1965. Individualized whistle contours in bottlenosed dolphins (*Tursiops truncatus*). *Nature*, 207 434-435.

[2] Bazúa-Durán, C., 2004. Differences in the whistle characteristics and repertoire of Bottlenose and Spinner Dolphins. *Anais da Academia Brasileira de Ciências*, 76(2) 386-392.

[3] Hoffmann, Franca, 2014. CCA mini-project report: *Dolphin Whistle Detector Algorithms: A New Robust Dolphin Detector using Coherence Enhanced Diffusion*

[4] Dobbins, P. and F. Hoffman, A new, robust, dolphin whistle detector algorithm, *Institute of Acoustics 40th Anniversary Conference*, NEC, Birmingham, 15-16 October 2014

Modelling the drug supply chain in clinical trials

Dr Paul D. Metcalfe (AstraZeneca)

Clinical trials are run across across many sites, and it is necessary to supply each site with the drugs that are needed for the patients that it will treat. It is important that whenever a patient turns up at a site we can supply them with the drugs they need; failure to do so is bad for the patient and can compromise the whole trial. The manufacture, transport, and storage of all these novel compounds is slow and difficult, so in order to make rapid progress and develop new medicines quickly and efficiently we want to only produce and store just enough drug, whilst controlling the risk of being unable to supply a patient with the drugs that they need.

For our purposes we can model the supply chain as a tree, with a factory as the root node. This may supply one or more depots, which may supply other depots or sites directly (and so on). So depots and sites each stock some level of drug, which is decreased as they supply patients or child depots and sites. But as the level of drug decreases they are free to order from their upstream.

Patients arrive at a site randomly (say, according to a Poisson process with known — or assumed — rate) and are then randomized to a treatment arm. Within each treatment arm patients may — randomly — undergo dose escalations or reductions that unexpectedly change their need for drug from that which is expected.

A number of questions may be explored within this project, for example:

1. Optimal strategy: How and when should sites and depots order from their upstream? Can anything analytical be done? We want to minimize the costs of manufacture, shipping, and storage, but also control the risk of being unable to supply a patient with the drugs that they need. Is there a good numerical approach to finding the optimum strategy?
2. Constrained supply, few patients, many sites: if patients that qualify for the trial are uncommon it's necessary to open many sites in order to recruit enough patients. Most sites won't recruit anyone; some will recruit many more than average. If supply is so constrained that it's not possible to start the trial with drug at every site how should we proceed?
3. High cost comparator: if one of the drugs in the trial has disproportionately high costs how does this affect supply chain strategies?
4. Anything else: anything that the candidate deems to be both useful and interesting.

Interested candidates should contact Dr Metcalfe to discuss the problem.

Optimal rational approximation and the double interval Zolotarev problem in the complex plane

Paul Childs
Schlumberger Gould Research, Cambridge, UK
childs4@slb.com

Acknowledgements: Vladimir Druskin, Leonid Knizhnerman and Stefan Guttel.

Problem statement

We consider the near best L^∞ approximation of $1/\sqrt{z}$ (or closely related functions) by a rational function, over certain curves in the complex plane.

Zolotarev's fourth problem [1], [6] provides the best rational approximation $r_{n-1,n}(z)$ of type $(n-1,n)$ to $\gamma = 1/\sqrt{z}$ over a real interval $I = [a,b]$, where $0 < a < b$ in the sense that

$$r_{n-1,n} = \underset{r_{n-1,n}}{\arg\min} \left\| \frac{\frac{1}{\sqrt{z}} - r_{n-1,n}(z)}{\frac{1}{\sqrt{z}} + r_{n-1,n}(z)} \right\|_{L^\infty(I)} . \tag{1}$$

Here, $r_{n-1,n}$ is a real rational function of type $(n-1,n)$ with a monic denominator, and I is the real interval $[a,b]$. Zolotarev's explicit solution involves Jacobi elliptic functions to determine the interpolation points. To make use of this in applications, we require the rational approximation $r_{n-1,n}(z)$ should be expressed as a Stieltjes continued fraction (or S-fraction) [2] with coefficients $\hat{h}_i$ $(0 \le i \le n-1)$ and h_i $(1 \le i \le n)$:

$$r_{n-1,n}(z) = \cfrac{1}{\hat{h}_0 z + \cfrac{1}{h_1 + \cfrac{1}{\hat{h}_1 z + \ldots \cfrac{1}{h_{n-1} + \cfrac{1}{\hat{h}_{n-1} z + \frac{1}{h_n}}}}}} . \tag{2}$$

The Zolotarev rational approximation to $1/\sqrt{z}$ over the real interval $[a,b] \in \mathbb{R}^+$, where $0 < a < b$, is readily extended to the real interval $[c,d] \in \mathbb{R}^-$ where $c < d < 0$. Recently [5] showed how to construct a nearly-optimal rational approximation to $1/\sqrt{z}$ on the union of two real intervals

$$I = [a,b] \cup [c,d] \text{ , where } c < d < 0 < a < b . \tag{3}$$

For practical use in numerical analysis (see Appendix), we actually require the best approximation to a modified square root $\gamma_1 = 1/\sqrt{z + \frac{z^2}{4}}$, rather than $\gamma = 1/\sqrt{z}$ over the union of *two* disjoint finite intervals in the *complex* plane:

$$I = I_{pr} \cup I_{ev} , \tag{4}$$

$$\Re(z) < 0 \text{ and } \quad \Im(z) \leq 0 \quad \forall z \in I_{pr} \ , \tag{5}$$
$$\Re(z) > 0 \text{ and } \quad \Im(z) \leq 0 \quad \forall z \in I_{ev} \ . \tag{6}$$

We use the same approach from [5] to construct a near-best rational approximation over the union of complex intervals

$$I_{pr} = [c, d] - i\epsilon \ , \tag{7}$$
$$I_{ev} = [a, b] - i\epsilon \ , \epsilon \in \mathbb{R}^+ \ . \tag{8}$$

The approach we use to construct a near best approximation, akin to (1), follows [5], modified for complex intervals:

- Construct the $2n_1$ Zolotarev interpolation points for $\sqrt{z}$ on the interval $[a/b, 1]$ and the $2n_2$ interpolation points for $\sqrt{z}$ on the interval $[-1, -d/c]$ separately. Here we require $n_1 + n_2 = n$.
- Scale these interpolation points to $[a, b]$ and $[c, d]$ and shift them by $-i\epsilon$.
- Calculate the unique complex rational interpolant $r_{n-1,n}(z)$ of type $(n-1, n)$ which interpolates $\gamma_1(z)$ at these $2n$ points.
- Calculate the coefficients of the S-fraction (2).

For practical reasons, the value of n used in the rational approximation $r_{n-1,n}(z)$ of type $(n-1, n)$ should be small, say $5 < n \leq 20$. We specify n and ask how to construct $r_{n-1,n}(z)$ which is the near best rational approximation to $\gamma_1(z)$ in order to minimize relative L^∞ error in the sense of (1). Since we solved a rational interpolation problem for $\gamma_1(z)$ rather than $\gamma(z)$, the Zolotarev points are no longer optimal, but are nearly optimal for minimising the relative L^∞ approximation error.

The proposed project has two aims:

- How well does the theory from [5], [2] for real intervals extend to complex intervals for approximating $1/\sqrt{z}$, and how good is the near-best approximation to $\gamma_1(z)$ in this case ?
- What is the best numerical method for obtaining the rational approximation of $\gamma_1(z)$ over complex intervals and to obtain its S-fraction representation ? Is the problem always well-posed ?

A major part of the project would be to review the literature: for example how far do the results from [3] provide an answer to this problem ? A related problem of practical importance is when I_{pr} and I_{ev} are not simply intervals shifted parallel to the real axis, but are smooth, non-intersecting, finite curves in the 3^{rd} and 4^{th} quadrants of the complex plane.

For a single interval $I = [a, b] \in \mathbb{R}^+$, it is known that all the coefficients $\hat{h}_i$ and h_i in (2) are real, positive and increasing if $r_{n-1,n}$ has negative simple poles and positive residues. Is it possible to say anything about the coefficients when there are two complex intervals ?

Appendix: practical application

This appendix has been added to motivate a practical application of near-best complex rational approximation. When constructing numerical boundary conditions for the wave equation, we require to approximate the so-called Neumann-to-Dirichlet (NtD) map on the boundary of a domain. Amazingly, in simple but representative cases, this can be reduced to the problem of finding the near best rational approximation of $1/\sqrt{z}$ which minimizes the relative L^∞ error, over certain intervals in the complex plane. Another practical application is in numerical linear algebra, when designing preconditioners for iterative solution of the Helmholtz equation [4]. The use of complex intervals I_{pr} and I_{ev} is required in this case.

The coefficients in (2) are used to construct numerical boundary conditions and represent the Neumann-to-Dirichlet map for the scalar wave equation for $u(\mathbf{x})$

$$k^2(\mathbf{x})u + \nabla^2 u = 0\,, k^2 \in \mathbb{C} \qquad , \Re(k^2) > 0 \quad , \Im(k^2) \geq 0\,. \tag{9}$$

For $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$, consider the half space $x_1 \leq 0$. Assuming k to be constant, and taking the Fourier transform in the x_2-direction with dual variable ξ_2 leads to an equation for $u(x_1, \xi_2)$:

$$k^2 u + \frac{\partial^2 u}{\partial {x_1}^2} - {\xi_2}^2 = 0\,, \tag{10}$$

which has fundamental solutions

$$u(x_1, \xi_2) = e^{\pm\sqrt{({\xi_2}^2 - k^2)x_1}}\,. \tag{11}$$

Well-posedness requires the solution to be finite as $x_1 \to -\infty$. Hence, for $\xi_2 \in \mathbb{R}$, the symbol of the Neumann-to-Dirichlet (NtD) operator at $x_1 = 0$, for given ξ_2, is $-1/\sqrt{{\xi_2}^2 - k^2}$. The practical application of the Zolotarev approximation to $1\sqrt{z}$ appears when (10) is discretized by a finite difference scheme - see [2], [5] and [4] for details. Consider the discrete form of (10),

$$(k^2 - {\xi_2}^2)u_j + \frac{u_{j+1} - 2u_j + u_{j-1}}{h^2} = 0\,, \tag{12}$$

for $-\infty < j \leq 0$, with mesh size h. The dispersion relation corresponding to the well-posed solution leads to a one-sided numerical approximation of the inverse of the NtD (i.e. the Dirichlet-to-Neumann map or DtN),

$$-\frac{u_0 - u_{-1}}{u_0} = -\frac{1}{2}z + \sqrt{z + \frac{z^2}{4}}\,, \tag{13}$$

where $z = -(k^2 - {\xi_2}^2)h^2$. This motivates the rational approximation for $\gamma_1(z)$. The S-fraction (2) corresponds to a finite-volume discretization of (10),

$$(k^2 - {\xi_2}^2)u_j + \frac{1}{\hat{h}_j}\left(\frac{u_{j+1} - u_j}{h_j} - \frac{u_j - u_{j-1}}{h_{j-1}}\right) = 0\,, \tag{14}$$

for $j = 1, \ldots, n-1$, and with the additional constraint $u_n = 0$. The coefficients h_j and $\hat{h}_j$ are step lengths on primary and dual grids - see [2] and [5] for details.

Only a very brief motivation for the problem has been given here in the interest of space. This relates the optimal numerical approximation of certain rational approximations in the complex plane to practical applications.

References

[1] N. I. Achieser: *Theory of Approximation*, Frederick Ungar, New York, 1956.

[2] Sergey Asvadurov, Vladimir Druskin, Murthy Guddati and Leonid Knizhnermann: *On optimal finite-difference approximation of PML*, SIAM J. Numer. Anal., 41(1), pp.287-305, 2003.

[3] B. Le Bailly and J. P. Thiran: *Optimal Rational Functions for the Generalized Zolotarev Problem in the Complex Plane*, SIAM J. Numer. Anal., 38(5), pp.1409-1424, 2000.

[4] P. N. Childs, V. Druskin, L. Knizhnerman: *Preconditioning the Helmholtz Equation using approximate Dirichlet-to-Neumann operators on optimal grids*, IMA Conference on Numerical Linear Algebra and Optimisation, University of Birmingham, 3rd-5th September 2014.

[5] Vladimir Druskin, Stefan Guttel and Leonid Knizhnerman: *Near-optimal perfectly matched layers for indefinite Helmholtz problems*, MIMS EPrint: 2013.53, University of Manchester, 2013.

[6] A. A. Gončar: *Zolotarev Poblems connected with rational functions*, Mat. Sbornik, Tom 78(210), no. 4, 1669.

Image Processing for Cultural Heritage Conservation

The topic of my external project involves the Cultural Heritage conservation of a particular fresco with tools from math and physics. The project is supervised by Dr Claudia Daffara of University of Verona (Italy) (she is a physicist at Department of Computer Science) and the CCA advisor is Dr Carola Schoenlieb. Also "Opificio delle Pietre Dure" of Florence is involved in this project in the person of Dr Paola Mariotti (it is the main public company devoted to art restoration in Italy).

The study I am going to carry on is possible thanks to the availability of FLIR System Inc. that lent us a thermal camera for this purpose. The dataset is processed with Matlab software (and it is likely that I will have a meeting with someone at Matlab in the next weeks).

I don't have an official abstract but I will try to summarize the project in an abstract-way:

The basic step of the project is to understand the physics behind the thermal image acquisition of FLIR camera and understand the principle of Thermal Quasi-Reflectography (TQR) technique recently introduced by Dr Claudia Daffara to unveil hidden layers and pentimenti in famous paintings. During the project, TQR is used to map the wall of a Leonardo da Vinci fresco at Castello Sforzesco in Milan, currently under restoration process in view of EXPO15. The restorers would like to see the hidden layers in order to restore the fresco in a proper way. The specific problem to be overcome is that the dataset presents some inhomogeneities of lights in TQR so the goal of this project is to provide a mosaic starting from each single TQR image thanks to an orthophoto and use a shadow removal algorithm in order to uniform the lights without losing the information. A test about the speed of the algorithm proposed is also carried out.