Optimization Research at SOL

A review of the last 30 years

Michael SAUNDERS

Systems Optimization Laboratory

Dept of Management Science and Engineering

Stanford University

Stanford, CA 94305-4026, USA

saunders@stanford.edu

Mathematisches Forschungsinstitut Oberwolfach

Schwarzwaldstr. 9–11, 77709 Oberwolfach-Walke, Germany

January 9–15, 2005

Abstract

QPSOL, LSSOL, NPSOL, QPOPT,
MINOS, SQOPT, SNOPT, Knossos, PDCO, SpaseLoc,
SYMMLQ, MINRES, LSQR, LUSOL

These are optimization packages developed at SOL during the last 30 years by Gill, Murray, Saunders, and Wright, with much input from coauthors Paige, Murtagh, Friedlander, and Jin.

We give a personal history of the development of these codes. For example, LUSOL is the engine for most of the large-scale solvers (MINOS, SQOPT, SNOPT, Knossos), as well as for ZIP, MILES, PATH, lp_solve.

Nonlinear Optimization

$$\begin{array}{ll} \underset{x \in R^n}{\text{minimize}} & \phi(x) \\ \text{subject to} & \ell \leq \begin{pmatrix} x \\ c(x) \\ Ax \end{pmatrix} \leq u \end{array}$$

$\phi(x)$	nonlinear objective function
$c(x)$	nonlinear constraint functions $\in R^m$
A	sparse matrix
ℓ, u	bounds

Assume functions are **smooth** with known **gradients**

MINOS

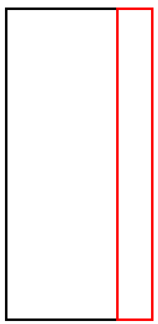
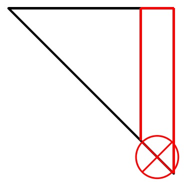
General sparse NLP

- 1975: Bruce Murtagh and MS
  NZ and SOL
- Sparse linear constraints, nonlinear obj
  Reduced-Gradient method (active set)
  LP + unconstrained optimization
  (Simplex + quasi-Newton)
- 1979: Sparse nonlinear constraints
  UNSW and SOL, extended Robinson 1972
- Assume functions and gradients are **cheap**
  GAMS and AMPL

QPSOL

Dense QP

- 1978: Gill & Murray
  1984: G, M, Saunders & Wright
- Dense linear constraints, indefinite QP obj $x^T H x$
- Orthogonal factors $A_k Q = (L \;\; 0), \quad Q = (Y \;\; Z)$
- Dense reduced Hessian $Z^T H Z = R^T R$

$Z =$  $R =$ 

- Only $\otimes$ needs care when Z, R get bigger

QPOPT

Dense LP and QP (G, M, & S 1995)

- Objective $= c^T x + \frac{1}{2} x^T G^T G x$
  or $\quad c^T x + \frac{1}{2} x^T H x$
- Positive definite, Semidefinite or Indefinite
- Dense linear constraints
- Later version of QPSOL (f77)

LSSOL

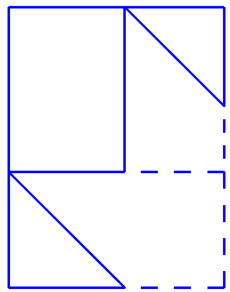
Dense Constrained Least Squares

$$\min \ \|Xx - b\|^2 \quad \text{s.t.} \quad l \leq \begin{pmatrix} x \\ Ax \end{pmatrix} \leq u$$

- 1971: Josef Stoer
- 1986: LSSOL: G, Hammarling, M, S & W
- Orthogonal factors

$$PXQ \ = $$

$$A_k Q \ = $$



- The only method that avoids forming $X^T X$
- We don't know how for sparse X

Lagrangians

NP	$\min\ \phi(x)$
	s.t. $\quad c(x) = 0$

Penalty Function

$$\min\ \phi(x) \qquad\qquad\qquad + \tfrac{1}{2}\rho_k \|c(x)\|^2$$

Augmented Lagrangian

$$\min\ \phi(x) - {y_k}^T c(x) + \tfrac{1}{2}\rho_k \|c(x)\|^2$$

Lagrangian in a Subspace

$$\min\ \phi(x) - {y_k}^T c(x) + \tfrac{1}{2}\rho_k \|c(x)\|^2$$

s.t. $\quad$ linearized constraints

LANCELOT

Conn, Gould & Toint 1992 and GALAHAD: G, Orban, T 2003

NP	$\min\ \phi(x)$	
	s.t. $c(x) = 0,$	$\ell \le x \le u$

BC_k	$\min\ \phi(x) - {y_k}^T c(x) + \frac{1}{2}\rho_k \|c(x)\|^2$	
	s.t.	$\ell \le x \le u$

BCL Method

- Solve subproblem BC_k
- Update dual variables: $y_{k+1} = y_k - \rho_k c(x_k^*)$
- Juggle ρ_k and repeat

NPSOL

Dense NLP (G, M, S & W 1986)

- Dense SQP method
  QP subproblems solved by LSSOL

- Search direction $(\Delta x, \Delta y)$
  QP$_k$ $\quad \min$ quadratic approx'n to Lagrangian
  $\quad$ s.t. linearized constraints

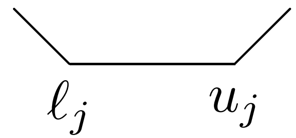
- Merit function
  Linesearch on augmented Lagrangian:

$$\min_{\alpha} \; L(x_k + \alpha\Delta x, \;\; y_k + \alpha\Delta y, \;\; \rho_k)$$

SQOPT

Sparse QP (G, M, & S 1997)

- Semidefinite QP obj $(x - x_0)^T H (x - x_0)$
- User routine computes Hx for given x
- Allows elastic bounds on variables and constraints



Application 1: Inside SNOPT

- $H = D + \sum v_j v_j^T - w_j w_j^T$ (Limited-memory quasi-Newton)

Application 1: Fused Lasso (Tibshirani, S et al. 2004)

$$\boxed{\min_{\beta} \|X\beta - y\|^2 \text{ s.t. } \sum |\beta_j| \le s_1 \text{ and } \sum |\beta_j - \beta_{j-1}| \le s_2}$$

- $Hx = X^T(Xx)$ (not ideal)
- Constraints on $\|\beta\|_1$ and $\|M\beta\|_1 \Rightarrow$ many $\beta_j = 0$ and $\beta_j = \beta_{j-1}$

SNOPT

Sparse NLP **(G, M, & S 2003; SIAM Review SIGEST 2005)**

- Sparse SQP method

  QP subproblems solved by SQOPT

- Search direction $(\Delta x, \Delta y)$

  QP_k $\quad \min$ limited-memory approx'n to Lagrangian

  $\qquad\quad$ s.t. $\quad$ linearized constraints

- Merit function

  Linesearch on augmented Lagrangian:

$$\min_{\alpha} \; L(x_k + \alpha\Delta x, \quad y_k + \alpha\Delta y, \quad \rho_k)$$

Infeasible Problems

or infeasible subproblems

SNOPT's solution – modify the original problem:

$$\min\ \phi(x) + \sigma \|c(x)\|_1$$

NP(σ)	$\min$	$\phi(x) + \sigma e^T(v + w)$	
	s.t.	$c(x) + v - w = 0,$	$v, w \geq 0$

Implemented by elastic bounds on QP slacks

SNOPT paper

revised for SIAM Review 2005

- LUSOL: Threshold Rook Pivoting for Basis Repair
- SYMMLQ on $Z^T H Z d = -Z^T g_{\mathrm{QP}}$ when many superbasics
- 1000 CUTEr and COPS 3.0 test problems
- Up to 40,000 constraints and variables
- Up to 20,000 superbasics (degrees of freedom)
- 900 problems solved successfully

Stabilizing
MINOS

Michael Friedlander's thesis 2002

Minimizing the Augmented Lagrangian

$$L(x, y, \rho) \;=\; \phi(x) - y^T c(x) + \tfrac{1}{2}\rho\|c(x)\|^2$$

BCL (LANCELOT)

$$\boxed{\begin{array}{ll} \displaystyle\operatorname*{minimize}_{x} & L(x, y_k, \rho_k) \\ \text{s.t.} & \\ & \ell \le x \le u \end{array}}$$

Globally convergent

LCL (MINOS)

$$\boxed{\begin{array}{ll} \displaystyle\operatorname*{minimize}_{x} & L(x, y_k, \rho_k) \\ \text{s.t.} & \text{linearized } c \\ & \ell \le x \le u \end{array}}$$

Quadratically convergent

An Elastic LC Subproblem

min aug Lagrangian

subj to bounds

$$\boxed{\begin{array}{ll} \underset{x}{\text{minimize}} & \phi(x) - {y_k}^T c(x) + \frac{1}{2}\rho_k \|c(x)\|^2 \\ \text{s.t.} & \ell \le x \le u \end{array}}$$

An Elastic LC Subproblem

min	aug Lagrangian	
subj to	linearized c	bounds

$$
\begin{array}{lll}
\underset{x}{\text{minimize}} & \phi(x) - {y_k}^T c(x) + \frac{1}{2}\rho_k \|c(x)\|^2 & \\
\text{s.t.} & \bar{c}_k(x) \qquad = 0, & \ell \le x \le u
\end{array}
$$

An Elastic LC Subproblem

min	aug Lagrangian + ℓ_1 penalty function
subj to	linearized c + elastic vars, bounds

$$\begin{array}{ll} \underset{x,v,w}{\text{minimize}} & \phi(x) - {y_k}^T c(x) + \frac{1}{2}\rho_k \|c(x)\|^2 + \sigma_k e^T(v+w) \\ \text{s.t.} & \bar{c}_k(x) + v - w = 0, \qquad v, w \geq 0, \quad \ell \leq x \leq u \end{array}$$

An Elastic LC Subproblem

min	aug Lagrangian + ℓ_1 penalty function
subj to	linearized c + elastic vars, bounds

$$\boxed{\begin{array}{ll} \underset{x,v,w}{\text{minimize}} & \phi(x) - {y_k}^T c(x) + \frac{1}{2}\rho_k \|c(x)\|^2 + \sigma_k e^T(v+w) \\ \text{s.t.} & \bar{c}_k(x) + v - w = 0, \qquad v, w \geq 0, \quad \ell \leq x \leq u \end{array}}$$

$$\Updownarrow$$

$$\boxed{\begin{array}{ll} \underset{x}{\text{minimize}} & \phi(x) - {y_k}^T c(x) + \frac{1}{2}\rho_k \|c(x)\|^2 + \sigma_k \|\bar{c}_k(x)\|_1 \\ \text{s.t.} & \qquad\qquad\qquad\qquad \ell \leq x \leq u \end{array}}$$

An Elastic LC Subproblem

min	aug Lagrangian + ℓ_1 penalty function
subj to	linearized c + elastic vars, bounds

$$\boxed{\begin{array}{ll} \underset{x,v,w}{\text{minimize}} & \phi(x) - {y_k}^T c(x) + \frac{1}{2}\rho_k \|c(x)\|^2 + \sigma_k e^T(v+w) \\ \text{s.t.} & \bar{c}_k(x) + v - w = 0, \qquad v, w \geq 0, \quad \ell \leq x \leq u \end{array}}$$

- Subproblems are always feasible
- Multiplier estimates are controlled by $\|\Delta y_k^*\|_\infty \leq \sigma_k$
- Eventually, x_k^* is same as for MINOS subproblem

The Stabilized LCL (sLCL) Method

1. Solve subproblem with optimality tol ω_k

$$(x_k^*, y_k^*) \leftarrow \text{Elastic LC subproblem } (x_k, y_k, \rho_k, \sigma_k)$$

2. If $\|c(x_k^*)\| \le \eta_k$

   Decrease ρ_k (finitely often)

   Reset σ_k (sufficiently large)

   Update $(x_{k+1}, y_{k+1}) \leftarrow (x_k^*, y_k^*)$

   else

   Reject (x_k^*, y_k^*)

   Increase ρ_k

   Decrease σ_k

- ρ_k and σ_k work together to ensure global convergence
- ω_k and η_k are adjusted dynamically to save work
- cf. Conn, Gould, Sartenaer, & Toint, 1996

KNOSSOS

Sparse NLP (Michael Friedlander & S 2004)

Fortran 90

- Dynamic memory
- LC solver “isolated”
- Threadsafe design (LC solver permitting)

LC subproblems

- Solve with MINOS or SNOPT
- Hot start on major itn 2, 3, 4, . . .
  (keep quasi-Newton Hessian or reduced Hessian)
- Early termination of subproblems

PDCO

Primal-Dual IPM for (Separable) Convex Opt

MATLAB (Saunders 1997–2004)

- Nominal problem

CO	$\min \ \phi(x)$
	s.t. $Ax = b, \qquad \ell \le x \le u$

- Regularized problem

COr	$\min \ \phi(x) + \frac{1}{2}\|D_1 x\|^2 + \frac{1}{2}\|r\|^2$
	s.t. $Ax + D_2 r = b, \qquad \ell \le x \le u$

$\nabla^2\phi(x),\ D_1 \succeq 0,\ D_2 \succ 0,$ all diagonal $\qquad A$ can be an operator

PDCO applications

- Basis Pursuit (Chen, Donoho, & S, 2001)
- Image reconstruction (Kim thesis, 2002)
- Maximum entropy (S & Tomlin, 2003)

PDCO on Web Traffic entropy problem

A is a 51000×662000 network matrix, nnz$(A) = 2$ million

Itn	mu	step	Pinf	Dinf	Cinf	Objective	center	atol	LSQR	Inexact
0			2.5	1.1	-6.7	-1.3403720e+01	1.0			
1	-5.0	0.267	2.4	1.1	-5.1	-1.3321172e+01	242.0	-3.0	5	0.001
2	-5.1	0.195	2.3	1.0	-5.3	-1.3220658e+01	36.9	-3.0	5	0.001
3	-5.2	0.431	2.1	0.9	-5.2	-1.2942743e+01	122.9	-3.0	5	0.001
4	-5.5	0.466	1.9	0.7	-5.3	-1.2711643e+01	41.8	-3.0	6	0.001
5	-5.7	0.671	1.4	0.2	-5.5	-1.2492935e+01	71.8	-3.0	9	0.001
6	-6.0	1.000	-0.0	-0.8	-5.8	-1.2367004e+01	2.7	-3.0	10	0.001
7	-6.0	1.000	-0.1	-2.3	-6.0	-1.2368200e+01	1.1	-3.0	9	0.002
8	-6.0	1.000	-1.1	-4.7	-6.0	-1.2367636e+01	1.0	-3.0	2	0.009
9	-6.0	1.000	-1.3	-5.7	-6.0	-1.2367655e+01	1.0	-3.0	7	0.015
10	-6.0	1.000	-2.5	-7.6	-6.0	-1.2367607e+01	1.0	-3.0	2	0.004
11	-6.0	1.000	-3.7	-8.6	-6.0	-1.2367609e+01	1.0	-3.5	8	0.004
12	-6.0	1.000	-5.9	-11.0	-6.0	-1.2367609e+01	1.0	-4.7	11	0.000

```
PDitns   =          12    LSQRitns =         79     time     =       101.4    (MATLAB)
                                                                      22.4    (C++)
```

SpaseLoc

Localization of Wireless Sensor Networks

MATLAB (Holly Jin + Carter, Saunders & Ye, 2004)

$$
\begin{array}{lllll}
\text{minimize} & \displaystyle\sum_{(i,j)\in N_1} |\alpha_{ij}| \;+ \sum_{(i,k)\in N_2} |\alpha_{ik}| & & & \\
\text{subject to} & \|x_i - x_j\|^2 - \alpha_{ij} & = & d_{ij}^2, & \forall\ (ni,j) \in N_1 \\
 & \|x_i - a_k\|^2 - \alpha_{ik} & = & d_{ik}^2, & \forall\ (i,k) \in N_2 \\
 & \|x_i - x_j\|^2 & \geq & r_{ij}^2, & \forall\ (i,j) \in \overline{N}_1, \\
 & \|x_i - a_k\|^2 & \geq & r_{ik}^2, & \forall\ (i,k) \in \overline{N}_2, \\
 & x_i,\ x_j \in \mathbf{R}^2, & & \alpha_{ij},\ \alpha_{ik} \in \mathbf{R}, & \\
 & i < j\text{: sensors} & & k\text{: anchors} &
\end{array}
$$

d_{ij} noisy distance data

a_k known positions of anchors

r_{ij} radio ranges

x_i sensors' positions (to be estimated)

Localization of Wireless Sensor Networks

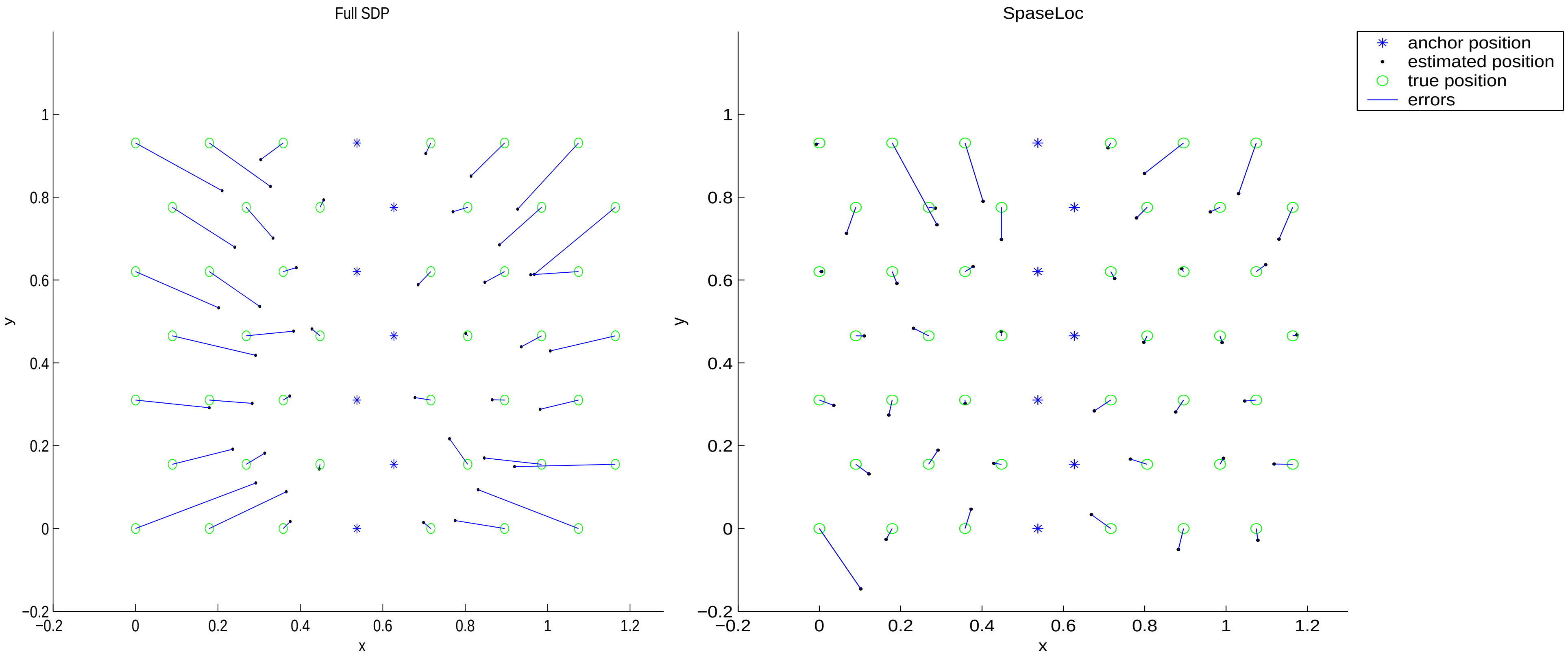
- Biswas and Ye (2003a): SDP relaxation — DSDP 2.0
  50 nodes: a few seconds
  200 nodes: too much time and storage

- Biswas and Ye (2003b): Parallel SDP subproblems — DSDP 2.0
  4000 nodes: 2 mins

- SpaseLoc (2004): Sequential SDP subproblems — DSDP 5.0
  4000 nodes: 25 secs

(DSDP = SDP solver of Benson and Ye)

Full SDP vs. SpaseLoc



Localization errors for full SDP model and SpaseLoc

LUSOL

Sparse LU + Updates (like Reid's LA05)

G, M, S & W 1987

Revised 1989–2004

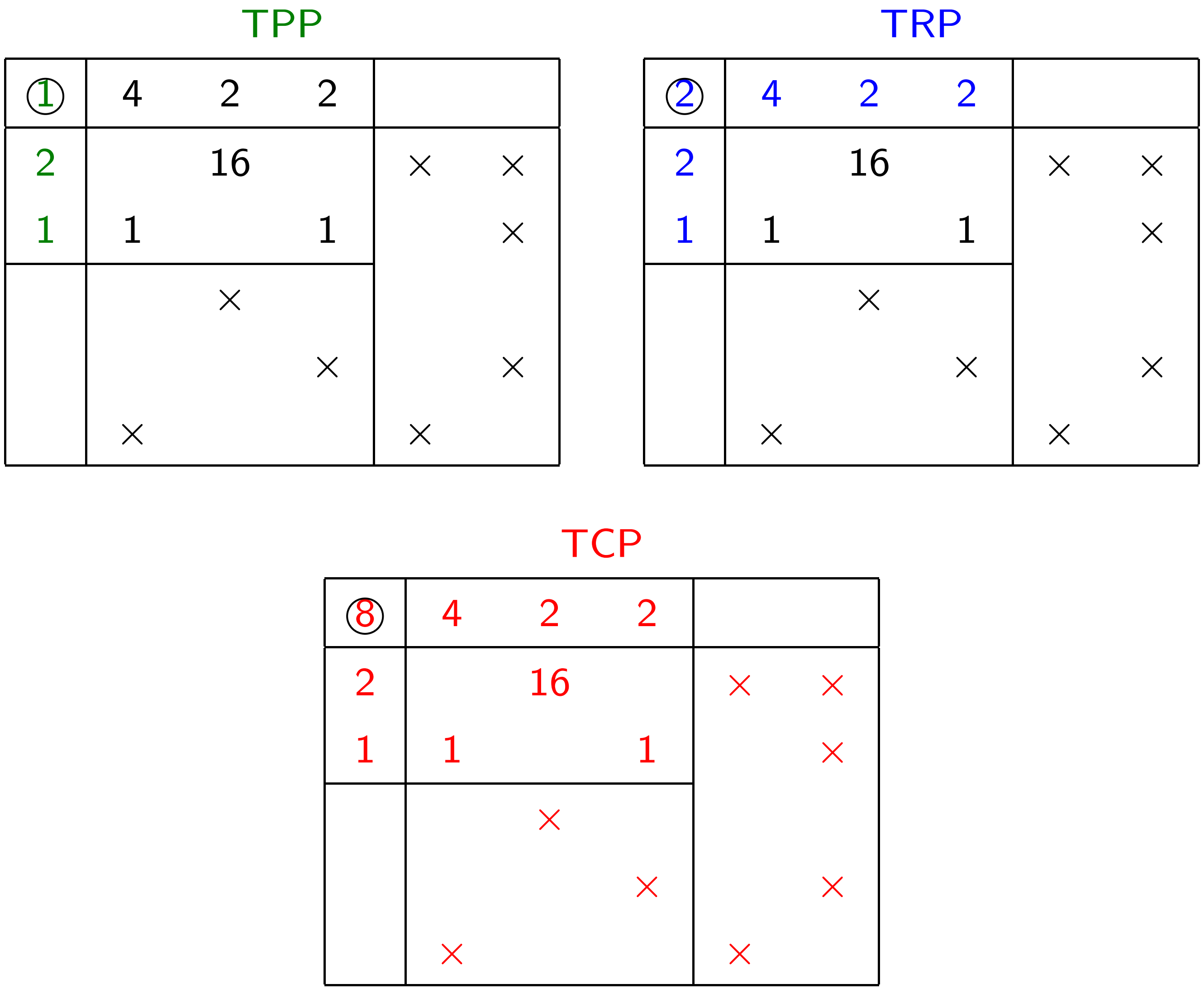
Part of Mike O'Sullivan's thesis 2002

- Square or rectangular $A = LU$
- Markowitz strategy for sparsity
- Threshold strategies for stability and detecting rank

TPP	Threshold Partial Pivoting	
TRP	Threshold Rook Pivoting	Rank-revealing
TCP	Threshold Complete Pivoting	Rank-revealing

- Bartels-Golub updating

Threshold Pivoting



Threshold Rook Pivoting is the right compromise for revealing singularity

LUSOL

BFP (Basis Factorization Package)

The engine for

MINOS
ZIP
MILES
PATH
SQOPT
Hence SNOPT and Knossos
lp_solve

Also Basis Repair

Square B — fix singularity
Rectangular $(B\ S)$ — find better B

Summary

Key concepts for Nonlinear Optimization

- Stable dense and sparse matrix factorizations
- Minimize augmented Lagrangian in (relaxed) subspace

Software

- MINOS, SNOPT in GAMS, AMPL, NEOS
- TOMLAB/SOL (Holmström) for MATLAB users
- SNOPT has its own MATLAB interface
- PDCO (MATLAB) $\min\ \phi(x)$ s.t. $Ax = b$ + bounds, A an operator
- SpaseLoc (MATLAB) Scalable sensor localization
- LUSOL (C version) now in open source system lp_solve
- LUSOL (original F77) will be in COIN-OR
- SYMMLQ, MINRES, LSQR (F77, MATLAB) CG-type iterative solvers

Other SOL Research

- Uday Shanbhag (with Walter Murray)
  Stochastic nonlinear programming, equilibrium programming, nonlinear facility location problems
- Che-Lin Su (with Dick Cottle)
  MPEC and EPEC (math programs with equilibrium constraints)
- Samantha Infeld (with Walter Murray)
  Trajectory optimization for spacecraft
- Yinyu Ye
  Large-scale (dual) SDP
  Wireless sensor network localization
  . . . !!
- http://www.stanford.edu/group/SOL/dissertations.html

The Lighter Side of Optimization

In New Zealand, the equivalent of the TV guide is called **The Listener**. Every week a "Life in New Zealand" column publishes clippings describing local events. The first sender of each accepted clipping receives a $5 Lotto Lucky Dip. The following clippings illustrate some characteristics of optimization problems in the real(?) world.

Robust solutions

> RECOVERY CARE gives you financial protection from specified sudden illness. You get cash if you live . . . and cash if you don't.

No objective function

> People have been marrying and bringing up children for centuries now.
>
> Nothing has ever come of it. (Evening Post, 1977)

Multiple objectives

"I had the choice of running over my team-mate or going onto the grass, so I ran over my team-mate then ran onto the grass", Rymer recalled later.

Obvious objective

He said the fee was increased from $5 to $20 because some people had complained it was not worth writing a cheque for $5.

Equilibrium condition

"The pedestrian count was not considered high enough to justify an overbridge", Helen Ritchie said. "And if there continues to be people knocked down on the crossing, the number of pedestrians will dwindle."

Constraints

ENTERTAINERS, DANCE BAND, etc. Vocalist wanted for New Wave rock band, must be able to sing.

DRIVING INSTRUCTOR Part-time position. No experience necessary.

HOUSE FOR REMOVAL in excellent order, $800. Do not disturb tenant.

Exactly one feasible solution

MATTHEWS RESTAURANT, open 365 nights. Including Mondays.

Buying your own business might mean working 24 hours a day. But at least when you're self-employed you can decide which 24.

Sergeant J Johnston said when Hall was stopped by a police patrol the defendant denied being the driver but after it was pointed out he was the only person in the car he admitted to being the driver.

His companion was in fact a transvestite, X, known variously as X or X.

Bound your variables

By the way, have you ever seen a bird transported without the use of a cage? If you don't use a cage it will fly away and maybe the same could happen to your cat. Mark my words, we have seen it happen.

Redundant constraints

If you are decorating before the baby is born, keep in mind that you may have a boy or a girl.

EAR PIERCING while you wait.

CONCURRENT TERM FOR BIGAMY (NZ Herald, 1990)

Infeasible constraints

I chose to cook myself to be quite sure what was going into the meals.

We apologize to Wellington listeners who may not be receiving this broadcast.

The model 200 is British all the way from its stylish roofline to its French-made Michelin tyres. (NZ Car Magazine)

$\geq$ or $\leq$?

BUY NOW! At $29.95 these jeans will not last long!

NOT TOO GOOD TO BE TRUE! We can sell your home for much less than you'd expect! (NZ Property Weekly)

The BA 146's landing at Hamilton airport was barely audible above airport background noise, which admittedly included a Boeing 737 idling in the foreground.

Yesterday Mr Palmer said "The Australian reports are not correct that I've seen, although I can't say that I've seen them".

It will be a chance for all women of this parish to get rid of anything that is not worth keeping but is too good to throw away. Don't forget to bring your husbands.

$\geq$ or $\leq$?

The French were often more blatant and more active, particularly prop X and number eight Y, but at least one All Black was seen getting his retaliation in first.

WHAT EVERY TEENAGER SHOULD KNOW — PARENTS ONLY

"Love Under 17" Persons under 18 not admitted.

"Keeping young people in the dark would not stop them having sex—in fact it usually had the opposite effect," she said.

NELSON, approximately 5 minutes from airport. Golf course adjacent. Sleeps seven all in single beds. Ideal for honeymoons. (Air NZ News, 1978)

Hard or soft constraints

The two have run their farm as equal partners for 10 years, with Jan in charge of grass management, Lindsay looking after fertilizer, and both working in the milk shed. “We used to have our staff meetings in bed. That got more difficult when we employed staff!” (NZ country paper)

Elastic constraints

The Stationary Engine Drivers Union is planning rolling stoppages.

When this happens there are set procedures to be followed and they are established procedures, provided they are followed.

APATHY RAMPANT? Not in Albany—the closing of the electoral rolls saw fully 103.49 percent of the area's eligible voters signed up.

Auckland City ratepayers are to be reminded that they can pay their rates after they die. (Auckland Herald, 1990)

He was remanded in custody to appear again on Tuesday if he is still in the country.

Convergence

"There is a trend to open libraries when people can use them", he says.

Mayor for 15 years, Sir Dove-Myer wants a final three years at the helm "to restore sanity and stability in the affairs of the city".

Applications

(Yachting) It is not particularly dangerous, as it only causes vomiting, hot and cold flushes, diarrhoea, muscle cramping, paralysis and sometimes death . . .

(Boating New Zealand, 1990)

(Ecological models) CAR POLLUTION SOARS IN CHRISTCHURCH—BUT CAUSE REMAINS MYSTERY

Nappies wanted for window cleaning. Must be used.

Stochastic optimization

Life is Uncertain.
Eat Dessert First. (Marie Callender T-shirt)

Always room for improvement

The owner Craig Andrew said the three main qualities for the job were speed, agility and driving skills. “Actually, Merv has none of those, but he’s still the best delivery boy we’ve had”, he said.

Integer Variables (0, 1, 2, . . .)

0 or 1 is sometimes not enough

When Taupo police arrested
a Bay of Plenty man
for driving over the limit
they discovered he was a bigamist

Nelson Mail, 5/04