Accurate and Efficient Solution of Linear and Nonlinear ME Models

Ding Ma, Michael Saunders, and Laurence Yang
Stanford University and University of California, San Diego

ICME INSTITUTE for COMPUTATIONAL & MATHEMATICAL ENGINEERING at STANFORD UNIVERSITY

UCSD

The need for quadruple precision

Carrying somewhat more precision in the arithmetic than twice the precision carried in the data and available for the result will vastly reduce embarrassment due to roundoff-induced anomalies

Default evaluation in Quad is the humane option

— *William Kahan (2012)*

Quad datatypes are now available

GCC provides `real(16)` and `float128` in gfortran and C, C++

This is the humane option for producing quad software

We use double-MINOS and quad-MINOS (f77 sparse LP/NLP) called from f77, f90, or python

Metabolic networks — flux balance analysis (FBA)

Constraint-based models enable the study of metabolism at genome-scale

- M models: multiscale reconstructions of metabolism
- ME models include protein expression (even more multiscale)
- Stoichiometric matrix $\boldsymbol{S}$, fluxes $\boldsymbol{v}$, growth rate μ
  Most coefficients are moderate: $S_{ij} = 0, \pm 1, \pm 2$
  Some coefficients are large: $S_{ij} = 10,000$
- Similarly for fluxes because of coupling constraints:
  $$v_i / v_j \ge \mu / k_{\mathit{eff}} \quad \Rightarrow \quad v_i \ge (\mu / k_{\mathit{eff}}) v_j$$

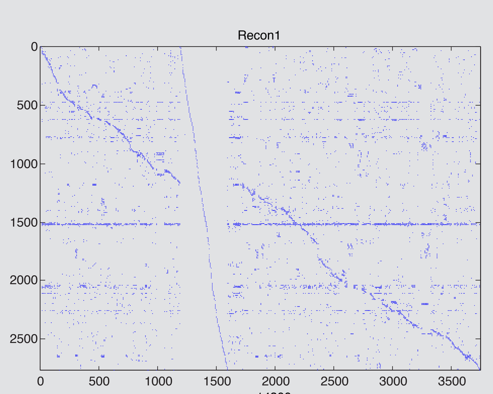
Models are linear when μ is fixed:

$$\begin{array}{ll} \max & c^T v \\ \text{st} & Sv = b \\ & \text{bounds on } v \end{array}$$

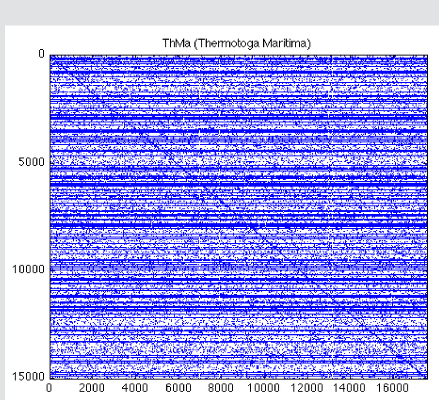
Solved by

- openCOBRA toolbox (CPLEX, glpk, Gurobi, Mosek, MINOS, . . .)
- MONGOOSE toolbox (exact simplex solver QSOpt ex)

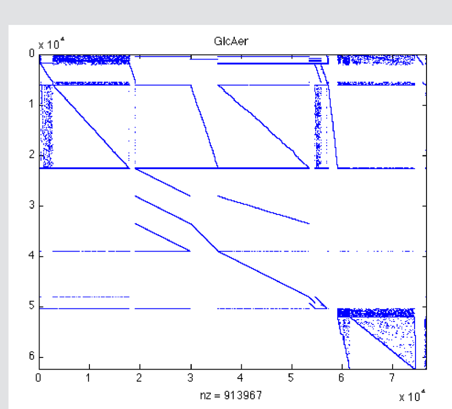
Sparse stoichiometric matrix S



Recon1



Thermotoga maritima



GlcAerWT

- Exact simplex as in MONGOOSE can handle nonlinear μ via binary search, but is not scalable to ME models
- Quad-precision LP and NLP provide a balance between reliability and speed

Linear ME model of *E. coli* — double-MINOS, quad-MINOS LP

Problem GlcAerWT (Thiele, Fleming, et al. 2012), **68300** $\times$ **76664**

Step 1: double-MINOS, cold start, scaling

```
Problem name                  GlcAerWT   EXIT -- the problem is infeasible
No. of iterations                62856   Objective value     -2.4489880182E+04
No. of infeasibilities              41   Sum of infeas        1.5279397622E+01
No. of degenerate steps          33214   Percentage                      52.84
Max x         (scaled)    68680 4.4E+06  Max pi        (scaled)    54979 1.4E+02
Max x                     62607 1.0E+09  Max pi                    25539 3.0E+02
Max Prim inf(scaled)     134382 6.5E+00  Max Dual inf(scaled)      70913 1.2E-05
Max Primal infeas        129844 1.0E+04  Max Dual infeas           23177 2.0E-05
Time for solving problem       9707.28 seconds
```

Step 2: quad-MINOS, warm start, scaling

```
Problem name                  GlcAerWT   EXIT -- optimal solution found
No. of iterations                 5580   Objective value     -7.0382449681E+05
No. of degenerate steps           4072   Percentage                      72.97
Max x         (scaled)    59440 3.7E+00  Max pi        (scaled)    40165 8.1E+11
Max x                     61436 6.3E+07  Max pi                    25539 2.4E+07
Max Prim inf(scaled)      83602 3.8E-16  Max Dual inf(scaled)      11436 4.4E-19
Max Primal infeas         83602 1.7E-07  Max Dual infeas           24941 8.6E-27
Time for solving problem       3995.58 seconds
```

Step 3: quad-MINOS, warm start, no scaling

```
Problem name                  GlcAerWT   EXIT -- optimal solution found
No. of iterations                    4   Objective value     -7.0382449681E+05
No. of degenerate steps              0   Percentage                       0.00
Max x                     61436 6.3E+07  Max pi                    25539 2.4E+07
Max Primal infeas        142960 1.3E-19  Max Dual infeas            6267 9.4E-22
Time for solving problem         60.07 seconds
```

Nonlinear ME model — variables μ, v, w

μ = growth rate

$$\begin{array}{lrl} \max & \mu & \\ \text{st} & \mu A v + B v & = 0 \\ & S v & = b \\ & \text{bounds on } v & \end{array} \qquad \equiv$$

A and B overlap

$$\begin{array}{lrl} \max & \mu & \\ \text{st} & \mu A v + w & = 0 \\ & B v - w & = 0 \\ & S v & = b \\ & \multicolumn{2}{l}{\text{bounds on } v\text{, none on } w} \end{array}$$

Nonlinear ME model — quad-MINOS NLP

tinyME (Yang et al. 2014), **2512** $\times$ **2828**

```
     ================================
     quadMINOS  5.6       (Nov 2014)
     ================================

Begin tinyME-NLP   cold start NLP with mu = mu0
Itn      304 -- linear constraints satisfied.
Calling funcon.  mu =  0.800000000000000000000000000000004
nnCon, nnJac, neJac        1073        1755        2681
funcon  sets    2681   out of    2681   constraint gradients.
funobj  sets       1   out of       1   objective  gradients.

Major minor  step     objective  Feasible Optimal   nsb   ncon penalty
    1   304T 0.0E+00  8.00000E-01 6.1E-03 2.1E+03     0      4 1.0E+02
    2   561T 1.0E+00  8.00000E-01 2.6E-14 3.2E-04     0     46 1.0E+02
    3    40T 1.0E+00  8.28869E-01 5.4E-05 3.6E-05     0     87 1.0E+02
    4      7 1.0E+00  8.46923E-01 1.2E-05 2.9E-06     0     96 1.0E+02
    5      0 1.0E+00  8.46948E-01 4.2E-10 2.6E-10     0     97 1.0E+02
    6      0 1.0E+00  8.46948E-01 7.9E-23 1.2E-20     0     98 1.0E+01

EXIT -- optimal solution found

No. of iterations                912   Objective value      8.4694810579E-01
No. of major iterations            6   Linear objective     0.0000000000E+00
Penalty parameter           1.000000   Nonlinear objective  8.4694810579E-01
No. of calls to funobj            98   No. of calls to funcon              98
No. of superbasics                 0   No. of basic nonlinears            786
Max x        (scaled)      12 5.6E-01  Max pi       (scaled)      103 8.3E+05
Max x                    1020 6.1E+01  Max pi                     103 9.7E+03
Max Prim inf(scaled)        0 0.0E+00  Max Dual inf(scaled)         9 2.9E-14
Max Primal infeas           0 0.0E+00  Max Dual infeas              9 1.3E-18
Nonlinear constraint violn    1.9E-20
```

Nonlinear ME model — quad-MINOS NLP

solveME (Yang et al. 2015), **11386** $\times$ **18755**

```
Calling minoss. Warm start with provided basis (hs)
Itn       32 -- linear constraints satisfied.
Calling funcon.  mu =  0.832815729997476367249118875820191994
nnCon, nnJac, neJac         2629        16126        14322
funcon  sets   14322   out of   14322   constraint gradients.
funobj  sets       1   out of       1   objective  gradients.

Major minor  step     objective  Feasible Optimal   nsb   ncon penalty BSwap
    1    32T 0.0E+00  8.32816E-01 4.3E-13 1.0E+03     0      4 1.0E+02     0

   19    40T 1.0E+00  8.32816E-01 2.5E-16 1.0E-03     0    743 1.0E+02     0
   20    40T 1.0E+00  8.32816E-01 1.0E-21 9.3E-04     0    784 1.0E+02     0

   23    40T 1.0E+00  8.55337E-01 3.4E-07 5.7E-05     0    907 1.0E+02     0
   24    40T 1.0E+00  8.55664E-01 2.1E-08 6.6E-07     0    948 1.0E+02     0
Itn      979 --       10  nonbasics set on bound, basics recomputed
   25     11 1.0E+00  8.55664E-01 7.0E-17 1.8E-11     0    961 1.0E+02     0
   26      0 1.0E+00  8.55664E-01 9.3E-19 8.0E-29     0    962 1.0E+01     0

EXIT -- optimal solution found

No. of iterations                979   Objective value      8.5566388920E-01
No. of major iterations           26   Linear objective     0.0000000000E+00
Penalty parameter           1.000000   Nonlinear objective  8.5566388920E-01
No. of calls to funobj           962   No. of calls to funcon             962
No. of superbasics                 0   No. of basic nonlinears           7896
Max x        (scaled)   12918 4.5E+01  Max pi       (scaled)     9674 3.3E+04
Max x                   14520 4.5E+01  Max pi                    4138 3.8E+03
Max Prim inf(scaled)        0 0.0E+00  Max Dual inf(scaled)         1 8.1E-24
Max Primal infeas           0 0.0E+00  Max Dual infeas           6827 1.1E-24
Nonlinear constraint violn    1.4E-19
```

References

- Murtagh and Saunders (1978, 1982)
  Large-scale linearly constrained optimization, *Math. Prog.* 14:41–72
  A projected Lagrangian algorithm and its implementation for sparse nonlinear constraints, *Math. Prog. Study* 23:349–352
- William Kahan (2011), Desperately needed remedies for the undebuggability of large floating-point computations in science and engineering,
  `http://www.eecs.berkeley.edu/~wkahan/Boulder.pdf`
- Thiele, Fleming, Que, Bordbar, Diep, and Palsson (2012)
  Multiscale modeling of metabolism and macromolecular synthesis in E. coli and its application to the evolution of codon usage, *PLOS ONE* 7(9), 18 pp
- Chindelevitch, Trigg, Regev, and Berger (2014)
  An exact arithmetic toolbox for a consistent and reproducible structural analysis of metabolic network models, *Nat. Commun.* 5(4893), 9 pp
- Ma and Saunders (2015)
  Solving multiscale linear programs using the simplex method in quadruple precision, Springer, to appear
- Yang, Ma, Ebrahim, Lloyd, Saunders, and Palsson (2015)
  solveME: fast and reliable solution of nonlinear ME models for metabolic engineering, *Metabolic Engineering*, submitted
- `stanford.edu/group/SOL/multiscale` `wwwen.uni.lu/lcsb`
  `opencobra.github.io/cobratoolbox` `mongoose.csail.mit.edu`

Funding

NIH U01GM102098
DOE ER65524

NATIONAL INSTITUTES OF HEALTH

DEPARTMENT OF ENERGY · UNITED STATES OF AMERICA

{saunders,dingma}@stanford.edu, laurence.yang@gmail.com