Accurate solution of FBA models using large-scale optimization in quadruple precision

Ding Ma, Michael Saunders, Ronan Fleming, and Ines Thiele
Stanford University and University of Luxembourg

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 (77 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 S, fluxes v, growth rate µ
▶ Most coefficients are moderate: Sij = 0, ±1, ±2
▶ Some coefficients are large: Sij = 10, 000
▶ Similarly for fluxes because of coupling constraints: vj/|vj| ≥ Keff 
▶ |vj| ≤ (µ/Keff) |vj|

Models are linear when µ is fixed:

• max cTv
• subject to Sv = b

Solved by:
• openCOBRA toolbox (CPLEX, glpk, Gurobi, MINOS, ...)
• MONGOOSE toolbox (exact simplex solver QSopt_ex)

Sparse stoichiometric matrix S

Recon1  Thermotoga maritima  GlcAerWT

▶ Exact simplex as in MONGOOSE is not scalable to the size of GlcAerWT, and cannot handle variable growth rate µ
▶ Quadr-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 x 76664

Step 1: double-MINOS, cold start, scaling

Problem name GlcAerWT EXIT — the problem is infeasible
No. of iterations 68256 Objective value -2.449880182E+04
No. of linear constraints 41 Subject of inf = 1.427593726E+05
No. of degenerate steps 33214 Percentage 52.84
Max x (scaled) 88680 1.0E+00 Max pi (scaled) 54973 1.0E+00
Max x 62237 1.0E+00 Max pi 25539 1.0E+00
Max Prim inf(scaled) 134392 6.5E+00 Max Dual inf(scaled) 70913 1.2E-05
Max Prim inf 83604 1.7E+00 Max Dual inf 23177 1.2E-05
Time for solving problem 9707.8 seconds

Step 2: quad-MINOS, warm start, scaling

Problem name GlcAerWT EXIT — optimal solution found
No. of iterations 4702 Objective value -7.038246918E-05
No. of linear constraints 59440 Percentage 72.97
Max x (scaled) 59440 3.7E-00 Max pi (scaled) 10156 8.1E+00
Max Prim inf(scaled) 14346 4.1E-19 Max Dual inf(scaled) 24914 4.6E-27
Time for solving problem 3995.6 seconds

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

Problem name GlcAerWT EXIT — optimal solution found
No. of iterations 4 Objective value -7.038246918E-05
No. of linear constraints 61436 Percentage 100
Max x 61436 3.6E-00 Max pi 25539 1.0E+00
Max Prim inf 83604 1.7E+00 Max Dual inf 23177 1.2E-05
Time for solving problem 60.1 seconds

Nonlinear ME model  quad-MINOS NLP

solveME (Yang et al., 2015), 11386 x 18755

Calling minos. Warm start with provided basis (bs)

fns 32 -- linear constraints satisfied.
funcon, mnu, mnc, mncA  2429 16126 14322
funcon sets 14322 out of 14322 constraint gradients, funcon sets 1 out of 1 objective gradients.
Major minor step objective Feasible Optimal nub nuc penalty (Eff)
1 327 0.000 0.82381610 1.30 -11 1.0e+00 0 4 1.0E-02 0
19 407 1.0E-03 0.82381610 2.5E-16 1.0E-03 0 743 1.0E-02 0
20 407 1.0E-03 0.82381610 1.0E-21 9.3E-04 0 784 1.0E-02 0
23 407 1.0E-03 0.55373670 2.1E-07 7.5E-05 0 907 1.0E-02 0
24 407 1.0E-03 0.55373670 2.1E-08 6.6E-07 0 948 1.0E-02 0
fns 999 -- 10 nondisps set on bound, bounds recomputed
11 1.0E-03 0.55373670 1.9E-19 8.2E-20 0 961 1.0E-01 0
16 0 1.0E-04 0.55373670 1.0E-19 8.2E-20 0 962 1.0E-01 0
EXIT — optimal solution found

No. of iterations 999 Objective value 8.55658902E-01
No. of major iterations 26 Linear objective 5.0000000000E-00
Penalty parameter 1.000000 Nonlinear objective 5.000000
No. of calls to funcon 962 No. of calls to funcon 962
No. of superbasics 0 No. of basic nonlinear 7896
Max x (scaled) 12354 5.4E+01 Max pi (scaled) 6914 3.3E+04
Max x 6914 3.3E+04 Max pi 14520 5.4E+01
Max Prim inf(scaled) 1.0E+00 Max Dual inf(scaled) 1.0E-10
Max Prim inf 1.0E+00 Max Dual inf 1.0E-10
Nonlinear constraint viol 1.4E-19

Nonlinear ME model

tinyME (Yang et al., 2014), 2512 x 2828

References
▶ Murtagh and Saunders (1978, 1982)
Large-scale linearly constrained optimization, Math. Prog. 14:41–72
▶ William Kahan (2011), Desperately needed remedies for the undefeasibility 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

NIH U01GM102098 DOE ER65524

Funding

2015 MSM Consortium Meeting, Bethesda, MD, Sep 8–10, 2015
{saunders,dingma}@stanford.edu, {ronan.mt.fleming,ines.thiele}@gmail.com
Special thanks to Laurence Yang, UCSD