Validated Numerical Optimization

George F. Corliss

Collaborators:

- Chenyi Hu, Houston
- R. Baker Kearfott, Southwestern Louisiana
- G. William Walster, Sun

Reference: Kearfott, Rigorous Global Search: Continuous Problems, Kluwer, 1966

Credits: Moore, Hansen, Neumaier, Caprani, Madsen, Walster, Ratz, Knüppel, Kearfott, ...

Applied?

"Consider ..."
- Pure math

"This could be used for ..."
- Applicable math

$n/2$ slides of science
1 slide of math or methods
$n/2$ slides of science
— APPLIED math

MU BIEN colloquium, January 18, 2002

Corliss Engine

Problem Domain

$f : X \subset R^n \to R$ continuous,
differentiable objective function

$$\min_{x \in X} f(x),$$

possibly with linear or nonlinear constraints.

Seek validated, tight bounds for minimizer X^*,
optimum value f^*.

Linear systems $Ax = b$ and nonlinear systems $G(x) = 0$

Challenges:

- Many local minima
- Flat objective
- High dimension

Applications

Sun: Applications keep the tool builders focused

MacNeal-Schwendler - Stress analysis of rocket nozzle designs, finite element analysis

Swiss Bank - Portfolio management: Foreign currency & options

Banc One - Portfolio management: Treasury bills/bonds

GE Medical Systems - Control of MRI signal, ODE parameter identification

Genome Theraputics - Gene space search, neural network

Computer algebra - Inexact GCD

Enabling Component: Intervals

Fundamental Theorem of Interval Analysis:
If a function is evaluated using interval arithmetic instead of floating-point arithmetic, the resulting interval is guaranteed to enclose the range of function values.

Global minimum can never be discarded

We consider in turn

1. Enabling components: interval analysis, automatic differentiation
2. Branch and bound for exhaustive search
3. Discarding where the solution is not
4. Validation: Guarantee the existence of a unique local minimum in some box
5. Enclosing a local minimum: Find a small box in which a unique local minimum is rigorously guaranteed to lie

Enabling Component: Intervals

$\boldsymbol{x} \circ \boldsymbol{y} := \{x \circ y : x \in \boldsymbol{x}, y \in \boldsymbol{y}\}$

Rosenbrock function $f(x,y) := 100 * \left(y - x^2\right)^2 + (1-x)^2$

$\boldsymbol{X} := (\boldsymbol{x}, \boldsymbol{y}) := ([0.7, 1], [0.0, 0.3])$

For $x \in \boldsymbol{x}$ and $y \in \boldsymbol{y}$,

$$
\begin{aligned}
f(x,y) &= 100 * \left(y - x^2\right)^2 + (1-x)^2 \\
&\in 100 * \left(\boldsymbol{y} - \boldsymbol{x}^2\right)^2 + (1-\boldsymbol{x})^2 \\
&= 100 * \left([0.0, 0.3] - [0.7, 1]^2\right)^2 + (1 - [0.7, 1])^2 \\
&= 100 * ([0.0, 0.3] - [0.49, 1])^2 + [0.0, 0.3]^2 \\
&= 100 * [-1.0, -0.19]^2 + [0.0, 0.09] \\
&= [3.61, 100.09].
\end{aligned}
$$

Guaranteed upper, lower bounds for range over $[0.7, 1] \times [0.0, 0.3]$

$f(0,0) = 1$ box $[0.7, 1] \times [0.0, 0.3]$ cannot contain a global minimum

Enabling Components
Automatic Differentiation

Compute required gradients, Hessians

Fast, accurate propagation of numerical values of derivatives

Not subject to the expression swell of purely symbolic differentiation

Not subject to the numerical inaccuracies of divided difference approximations.

Interval arithmetic bounds derivatives, too

$$x := \boldsymbol{x} = [0.7, 1.0]; \quad \nabla x := \left(\frac{\partial x}{\partial x}, \frac{\partial y}{\partial x}\right) = (1, 0)$$

$$y := \boldsymbol{y} = [0.0, 0.3]; \quad \nabla y := \left(\frac{\partial x}{\partial y}, \frac{\partial y}{\partial y}\right) = (0, 1)$$

evaluate the code list using interval arithmetic

$$f = [3.61, 100.09]; \nabla f = ([52.6, 400], [-200, -38])$$

0 is not in either interval $[52.6, 400]$ or $[-200, -38]$

f has no critical point in the box $[0.7, 1] \times [0.0, 0.3]$

- More at Dr. Feng's Computational Intelligence seminar, Thursday, October 13, 2-3:00

Branch and Bound

Maintain three lists of boxes:

$\mathcal{L}$ not fully analyzed;
$\mathcal{R}$ small boxes validated to enclose
a unique local minimum; and
$\mathcal{U}$ small boxes not validated either to enclose
a minimum or *not* to enclose a minimum.

To find enclosures for f^* and X^* such that

$$f^* = f(X^*) \leq f(X),$$

for all $X \in \boldsymbol{X}_0 \subset R^n$:

Branch and Bound

Initialize $\mathcal{L} = \{\boldsymbol{X}_0\}$ // List of pending boxes
$\overline{f} = +\infty$ // $f^* \leq \overline{f}$
DO WHILE $\mathcal{L}$ is not empty
    Get current box $\boldsymbol{X}^{(c)}$ from $\mathcal{L}$
    IF $\overline{f} < \boldsymbol{f}(\boldsymbol{X}^{(c)})$ THEN
        // We can guarantee there is
            no global minimum in $\boldsymbol{X}^{(c)}$
        Discard $\boldsymbol{X}^{(c)}$
    ELSE
        Attempt to improve $\overline{f}$
    END IF
    IF we can otherwise guarantee that there
        is no global minimum in $\boldsymbol{X}^{(c)}$ THEN
        Discard $\boldsymbol{X}^{(c)}$
    ELSE IF we can validate the existence of
        a unique local minimum in $\boldsymbol{X}^{(c)}$ THEN
        Shrink $\boldsymbol{X}^{(c)}$, and add it to $\mathcal{R}$
    ELSE IF $\boldsymbol{X}^{(c)}$ is small THEN
        Add $\boldsymbol{X}^{(c)}$ to $\mathcal{U}$
    ELSE
        Split $\boldsymbol{X}^{(c)}$ into sub-boxes, and add them to $\mathcal{L}$
    END IF
END DO

Discarding where the solution is NOT

Range on a sub-box $>$ a known upper bound for f^*

Non-zero gradient

Hessian of f is not positive definite

Fritz-John or Kuhn-Tucker conditions

Feasibility of constraints

Validation

Neumaier: Purposes of validated scientific computation:

- Obtain rigorous verification of existence and uniqueness, and
- Find guaranteed enclosures with narrow bounds.

Three separate issues:

- validation
- guaranteed enclosures
- tight enclosures

For validation, **Banach Fixed-Point Theorem:**
Let $T : X \to X$ be defined on a complete metric space X with metric ρ. Let α satisfy $0 < \alpha < 1$, and let

$$\rho(T(x), T(y)) \leq \alpha\rho(x, y)$$

for all x and $y \in X$.
Then T has a unique fixed point $x^* \in X$.

Validation
Interval Newton Method

Univariate interval Newton method

$$\boldsymbol{N}(f;\boldsymbol{x},\check{x}) := \check{x} - f(\check{x})/\boldsymbol{f}'(\boldsymbol{x})$$
$$\boldsymbol{x}^{(k+1)} := \boldsymbol{x}^{(k)} \cap \boldsymbol{N}(f;\boldsymbol{x}^{(k)},\check{x}^{(k)})$$

Multi-dimensional version:
$G := \nabla f$; $\check{X}$ approximate minimizer; $||G(\check{X})||$ is small
$\boldsymbol{A}$ enclosure of Jacobian $\boldsymbol{G}'(\boldsymbol{X})$
($=$ the Hessian of f, computed by automatic differentiation).

DO
- $\boldsymbol{A} := \boldsymbol{G}'(\boldsymbol{X})$
- Bound solution set of $\boldsymbol{A}(X - \check{X}) = -\boldsymbol{G}(\check{X})$
  - for $X - \check{X} \subset \boldsymbol{X1}$
- IF $\boldsymbol{X1} + \check{X} \subset$ Interior $(\boldsymbol{X1})$ THEN
  - Have validated existence and uniqueness
  - Done
- ELSE
  - $\boldsymbol{X} := \boldsymbol{X} \cap \left(\boldsymbol{X1} + \check{X}\right)$
  - $\check{X} :=$ midpoint $(\boldsymbol{X})$
- END IF

END DO

Box $\boldsymbol{X}$?

Where does the box $\boldsymbol{X}$ in which validation is attempted come from?

- Approximate optimizer $\check{X}$
- Verify existence in as small a box as possible
- Verify uniqueness in as large a box as possible

Enclosing the Solution

Runs until original box $\boldsymbol{X}$ is separated:

list $\mathcal{R}$	small boxes validated to enclose a unique local minimum; and
list $\mathcal{U}$	small boxes not validated either to enclose a minimum or NOT to enclose a minimum
discarded	regions where the global optimizer X^* cannot occur

or until time or memory resource limits are reached

GlobSol: Rigorous Global Solutions

For
- Global optimization
- Nonlinear systems

$f : X \subset R^n \to R$ continuous,
differentiable objective function

$$\min_{x \in X} f(x),$$

possibly with linear or nonlinear constraints.

Seek validated, tight bounds for minimizer X^*, optimum value f^*.

Uses
- Interval arithmetic
- Automatic differentiation

Algorithm components:
- Branch and bound
- Validated range bounds, objective & constraints
- Discard regions where solution is not
- Variants of interval Newton
- Constraint propagation

References:
- Kearfott, Rigorous Global Search: Continuous Problems
- `www.mscs.mu.edu/~globsol/`

MacNeal-Schwendler

Stanton, Hart (MacNeal-Schwendler); Fritz, Fisher, Johnson, Kearfott, Corliss
www.mscs.mu.edu/$\sim$georgec/Pubs/students.html#Fritz

Stress analysis of rocket nozzle designs
finite element analysis
Low-dimensional design optimization

Traditional FEA:

$F = K(d) \cdot u$, where

$$\begin{array}{lcl} K(d) & = & n \times n \text{ elastic matrix (lb/in)} \\ F & = & n\text{-dimensional force vector (lb)} \\ \cdot(d) & & \text{represents dependence on design variables} \end{array}$$

Strain $= B(d) \cdot u$

Stress $=$ Strain $\cdot D$

To optimize something

- Adjust design variables d
- Resolve FEA equations

"All at Once" Method

Independent variables:

- Design variables
- Element displacements

Steps:

- Formulate a Finite Element Model
- Constraints: FEA equations
- Let Max strain $\geq$ strain(x)
- Minimize Max strain

Example. Objective: Max internal F

Equality constraints: $Ku - F = 0$

Inequality constraints: stress - stress limit ≤ 0

E.g.: 10 elements,

- F in [2999.69, 3000.30]
- 16 CPU sec (SPARC 20)
- using GlobSol

Swiss National Bank

Vignes (Swiss Bank); Ranjit, Daniels, Kearfott, Corliss
www.mscs.mu.edu/$\sim$georgec/Pubs/students.html#Ranjit

Modeling international currency trading
Holdings in 40 currencies, options
Options: call/put, long/short
Adjust each afternoon

Minimize maximum Value at Risk (VaR): How much might you loose?

$$\sigma_v^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \xi_{i,j} F_{i,j} \rho_{i,j} \sigma_{i,j},$$

Prices assumed bounded in historical ranges

Find Maximum possible loss
or Minimum variance
subject to constraints

For one two week period in October, 1997, the minimum variance portfolio of Swiss Frank, Mexican Peso, South Korean WON, Australian Dollar, Netherlands' Guilder, and Singapore Dollar had 74.7% Mexican Peso, 8.9% Australian Dollar, 8.4% Netherlands' Guilder, and 8.0% Singapore Dollar.

Swiss Bank - Call

Right, but not the obligation, to **buy** a sum of foreign exchange at a fixed exchange rate (strike price E) at a specified expiration date

European option: **at** the specified expiration date
American option: **on or before** expiration date

Values from Garman-Kohlhagen (G-K) model (derived from Black-Scholes) for European style foreign exchange options:

$$V_C = Se^{-r_f t} N(h) - Ee^{-r_d t} N\left(h - v\sqrt{t}\right)$$

where

S	-	spot rate, domestic to foreign currency units
E	-	exercise or strike price of the option
r_d	-	risk-free return in the domestic currency
r_f	-	risk-free return in the foreign currency
N	-	cumulative normal distribution
v	-	annual volatility
t	-	time remaining to expiration of the option, in years
h	$=$	$\left(\ln\left(\frac{S}{E}\right) + \left(r_d - r_f + \frac{v^2}{2}\right)t\right)\frac{1}{v\sqrt{t}}$

Swiss Bank - Call

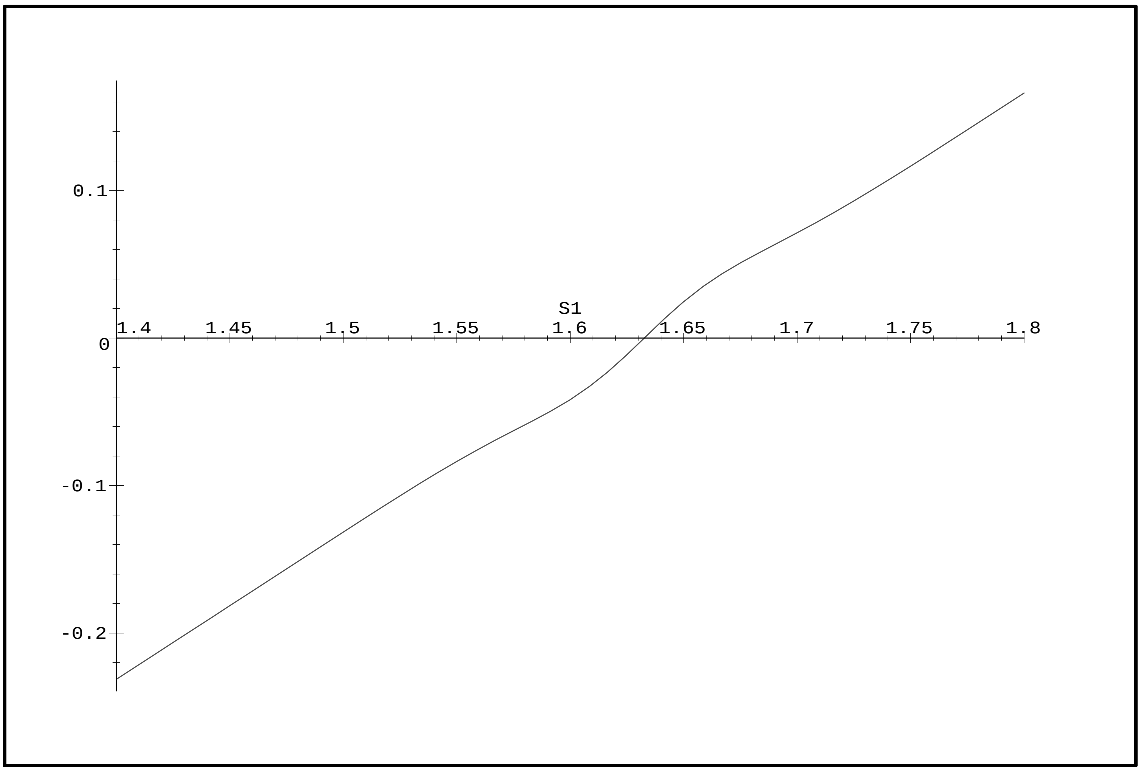

Figure 1: Call value V_C vs. strike price E

Swiss Bank - Put

Put: Right, but not the obligation, to **sell** a sum of foreign exchange at a fixed exchange rate (strike price E) at a specified expiration date

$$V_P = -Se^{-r_f t} N(-h) + Ee^{-r_d t} N\left(v\sqrt{t} - h\right)$$

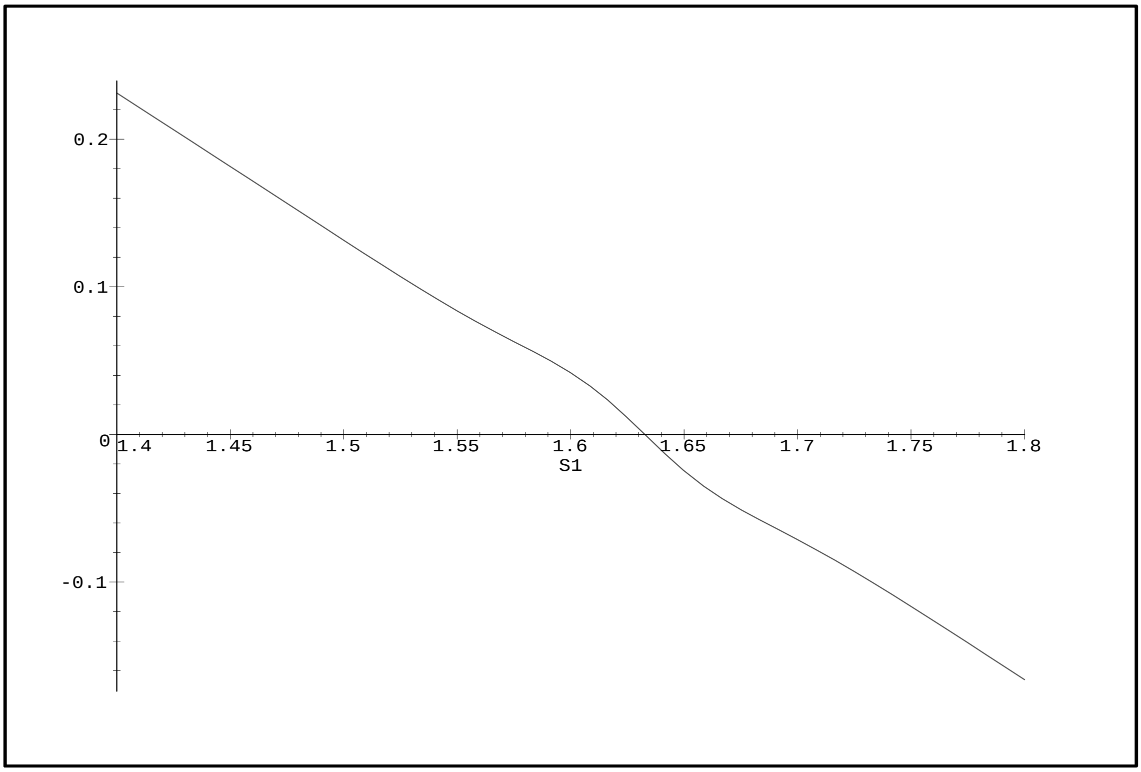

Figure 2: Put value V_P vs. strike price E

Swiss Bank - Objectives

Portfolio holds many contracts

- Which foreign currency?
- Face value?
- Strike price?
- Expiration?

Risks:

- Currency price?
- Interest rates?

Optimize

- Return or profit?
- Risk, uncertainty, or variance?
- Value at Risk (VaR)?

Swiss Bank - Value at Risk

VaR: Given historical price and interest rate fluxuations (99% confidence), what is the **most we might loose** from adverse market movement?

$$
\begin{aligned}
V_{Ci} &= S_i e^{-r_f t} N(h) - E e^{-r_d t} N\left(h - v\sqrt{t}\right) \\
V_{Pi} &= -S_i e^{-r_f t} N(-h) + E e^{-r_d t} N\left(v\sqrt{t} - h\right)
\end{aligned}
$$

$$\min\left(\sum_{i=1}^{nc} m_{Ci} V_{Ci} + \sum_{i=1}^{np} m_{Pi} V_{Pi}\right)$$

such that

$$S_{i0} - \beta\sigma_i \leq S_i \leq S_{i0} + \beta\sigma_i$$

Swiss Bank - Example

E.g. (in US $, domestic interest rate 5.56%)

Amount	Currency	Strike	Interest	Mature	Volit.
Call 100	British Pounds	1.64	6.87%	30 d.	12%
Put 100	British Pounds	1.64	6.87%	30 d.	12%
Call 100	German Marks	0.535	3.09%	30 d.	14%
Put 100	German Marks	0.535	3.09%	30 d.	14%

At 1.633 USD/BP and 0.545 USD/DM, current value is $6.75 (USD)

IF prices stay within their 99% confidence interval (2.33 historical standard deviations)

Minimum value = $6.19 (USD) at 1.641 USD/BP and 0.5338 USD/DM

Maximum loss = VaR = $6.74 - $6.19 = $0.56, or about 8.3% of current market value

If that is too much risk, seek additional hedging positions

Bank One

Myszka (Banc One); Thalacker, Julien, Daniels, Toumanoff, Corliss

Portfolio management using global optimization
Portfolio of Treasury bills, bonds

Maximize expected return

$$RR = -4 \sum_{i=1}^{n} P_i X_i (P_i - 1)/i$$

Constraints:

- Total value
- Cash flow
- Average duration
- Variance

Low dimension
Nearly flat objective
Uncertain parameters

For a typical period, GlobSol found the optimal allocation to be 9.7% in 3-month, 59.8% in 6-month, and 30.5% in 12-month T-bills, for an average rate of return of 5.398% per year. The solution required about 17 CPU seconds on a Sun SPARC 10.

Potential benefits: $\geq$ \$100K per DAY
Safe from Boss + Excel?

GE Medical Systems

Ganin (GEM); Yan, Yang, Zhu, Feng, Corliss

Control of magnetic resonance imaging signal to compensate for eddy currents

Fit output signal $G(t)$ by solution to DE

$$x' = f(x, t, a, u(t))$$

parameters a, control function $u(t)$

Minimize $\|G(t) - x(t)\|_2$

Given u_{input} and G_{observed}
Formulate model f & optimal parameters a

Given model f & parameters a
Construct optimal control $u(t)$

Strategy ("All-together", Dennis & Williamson):

- Collocate with piecewise cubics
- Constraint: Continuous at nodes
- Constraint: Differentiable at nodes
- Constraint: "Satisfies" the DE
- Solve constrained optimization

Genome Theraputics

Wang, Tonellato, Corliss

Integration of results of gene search queries processed by several different search engines

Present gene sequence U to different engines
They each return segments recognized
What is our best guess?

Neural network

$$\text{Node: } f\left(\sum_{i=1}^{n} a_i x_i\right)$$

Optimization is training NN

Lesson: Most neural networks are over-parameterized

Leads to singular optimization problem

Computer algebra

Chin, Corless, Corliss

Greatest common denominator of polynomials with floating-point coefficients

$\epsilon \geq 0$ GCD of polynomials p, q is a polynomial $d(x)$ of highest possible degree such that $d(x)$ exactly divides

$$p(x) + \Delta p(x) \text{ and } q(x) + \Delta q(x),$$

where $\|\Delta p(x)\| \leq \epsilon, \|\Delta q(x)\| \leq \epsilon$

Minimize $\|\Delta p(x)\| + \|\Delta q(x)\|$

subject to

$$d(x)u(x) = p(x) + \Delta p(x)$$
$$d(x)u(x) = q(x) + \Delta q(x)$$

ODE Control with Switching?

Examples:

- Control Circuit for an Artificial Heart (Brown and Corliss)
- Water reservoir level control (Nedialkov)
- Stability of particle accelerators (Berz)

With Berz and Nedialkov

System described by system of ODEs:

- Multiple states
- Transitions due to switching functions

Challenges:

- Trajectories are slightly uncertain - tubes
- Switching times are slightly uncertain

Goals:

- Reliable computation of trajectories
- Guaranteed bounds on long-term behavior

Your problem?

When should I consider this technology?

- Mission-critical problems or subproblems
- Feasibility
- Many local critical points
- Singular or near-singular solutions
- Uncertain parameters
- Bounded observation errors
- Model validation
- Speed

Limitations

- High dimensions
- Speed
- Recoding