Examples of Running Constrained Optimization Codes

CEE 201L. Uncertainty, Design, and Optimization
Department of Civil and Environmental Engineering
Duke University
Henri P Gavin
Spring 2024

Objective function; constraint inequalities

The conventional form of a design optimizations is

$$\min_{x_1, x_2, \ldots, x_n} J = f(x_1, x_2, \ldots, x_n) \text{ such that}$$

$$g_1(x_1, x_2, \ldots, x_n) \leq 0, \quad x_{lb,1} \leq x_1 \leq x_{ub,1}$$

$$\vdots$$

$$g_m(x_1, x_2, \ldots, x_n) \leq 0, \quad x_{lb,n} \leq x_n \leq x_{ub,n}$$

where $[x_1, \ldots, x_n]$ is a vector of design variables, $f(x_1, \ldots, x_n)$ is the objective function to be minimized, $g_i(x_1, \ldots, x_n) \leq 0$ is the $i$-th out of $m$ constraint inequalities to be satisfied, and $x_{lb,i} \leq x_i \leq x_{ub,i}$ provide lower bounds (lb) and upper bounds (ub) on each of the design variables individually.

Convention for inequality constraints and constraint scaling

By convention, all constraints are expressed as an inequality compared to zero. And by convention, positive constraint values are “not ok”. A standard (conventional) constraint function $g(x)$ for an inequality

$$p(x) \leq q(x)$$

is an inequality compared to zero

$$g(x) = p(x) - q(x) \leq 0$$

If the numerical values of one constraint equation are much much larger or smaller than numerical values of other constraint equations, it is helpful to scale the constraints so that they all have values roughly around the range (-1,1). To do so ...

If $p$ is a positive constant, use $g(x) = 1 - q(x)/p \leq 0$. In code this is written $g = 1 - q/p$;
If $q$ is a positive constant, use $g(x) = p(x)/q - 1 \leq 0$. In code this is written $g = p/q - 1$;
If $p$ is a negative constant, use $g(x) = q(x)/p - 1 \leq 0$. In code this is written $g = q/p - 1$;
If $q$ is a negative constant, use $g(x) = 1 - p(x)/q \leq 0$. In code this is written $g = 1 - p/q$;
Example optimization problem

This document demonstrates the use of numerical methods to solve the following optimization problem:

$$\min_{x_1, x_2} J = f(x_1, x_2) = (x_1 - c_1)^2 + (x_2 - c_2)^2 + c_3 N$$

such that:

$$g_1(x_1, x_2) = a_1((x_1 - a_2)^2 + (x_2 - a_3)^2) - a_4 \leq 0,$$

$$g_2(x_1, x_2) = b_1((x_1 - b_2)^2 + (x_2 - b_3)^2) - b_4 \leq 0,$$

and

$$0 \leq x_i \leq 1 \quad i \in (1, 2)$$

where \([a_1, a_2, a_3, a_4], [b_1, b_2, b_3, b_4], \) and \([c_1, c_2, c_3], \) are numerical constants and \(N\) is an uncertain value modeled by a standard-normal distribution. The coefficient \(c_3\) determines the level of uncertainty in the objective. In this example, the unconstrained minimum of \(f(x)\) is at \((x_1^*, x_2^*) = (c_1, c_2).\)

Figure 1. Surfaces and contours of the objective, \(f(x)\) (red), and the constraints \(g_1(x)\) (green) and \(g_2(x)\) (blue). \(a = [1.0, 0.2, 0.5, 0.3], b = [-2.0, -0.5, 0.5, -1.5]\) and \(c = [0.8, 0.2, 0.0]\) (max headroom)
Examples of Running Constrained Optimization Codes

Code (with comments)

```matlab
function [ f , g ] = optim_example_analysis( x , constant )

% [ f , g ] = optim_example_analysis( x , constant )

% INPUT DESCRIPTION DIMENSION
% x a vector of design variables n x 1
% constant a cell array or vector of problem constants

% OUTPUT DESCRIPTION DIMENSION
% f the objective function corresponding to x 1 x 1
% g the constraint vector corresponding to x m x 1

% example of how to set up an optimization-oriented analysis
% H.P. Gavin, Dept. Civil & Environ. Eng'g, Duke Univ., 2011-1-25

x1 = x(1); % description of design variable number 1, units
x2 = x(2); % description of design variable number 2, units

a = constant{1}; % description of constant number 1, units
b = constant{2}; % description of constant number 2, units
c = constant{3}; % description of constant number 3, units

f = ( x1 - c(1) )^2 + ( x2 - c(2) )^2 + c(3)*randn; % the objective function

% optim_example_opt.m
% script for carrying out a numerical optimization
% H.P. Gavin, Dept. Civil & Environ. Eng'g, Duke Univ., 2011-1-25

x_lb = [ 0 ; 0 ]; % lower bound on design variables
x_ub = [ 1 ; 1 ]; % upper bound on design variables

x_init = x_lb + rand(2,1).*(x_ub - x_lb); % random initial guess, or ...
% a set of constants used in the analysis
a = [ 1.0 , 0.2 , 0.5 , 0.3 ];
b = [ -2.0 , -0.5 , 0.5 , -1.5 ];
c = [ 0.8 , 0.2 , 0.0 ];
constant = { a , b , c };

% algorithmic parameters ...
% msglevel tolX tolF tolG MaxEvals Penalty Exponent mMax errF
options = [ 3 0.001 0.001 0.001 5e2 0.5 0.5 50 0.01 ];

% ... ORSopt or NMAopt or SQPopt ...
[x_opt , f_opt , g_opt , cvg_hst] = NMAopt( 'optim_example_analysis' , x_init , x_lb , x_ub , options , constant );

plot_cvg_hst( cvg_hst , x_opt , 100); % plot the convergence history
```

To try these optimization solutions yourself, download these three optimization functions:

- Optimized Step-Size Random Search ORSopt
- Nelder-Mead with inequality constraints NMAopt
- Sequential Quadratic Programming SQPopt

and the following five helper functions:

- set values for optimization algorithm parameters optim_options.m
- average and coefficient of variation of augmented objective avg_cov_func.m
- constrain a vector to a box of specified size box_constraint.m
- plot the objective as a surface over two parameter values plot_surface.m
- plot the convergence history of variables and max constraint plot_cvg_hst.m

These .m-files and more may be found at http://www.duke.edu/~hpgavin/m-files
ORSopt Results

\[
\text{objective } f_{aug}(x_1, x_2) = 1.0154 \times 10^0 \\
\text{max}(g_{opt}) = -6.8954 \times 10^{-5}
\]

- Objective function convergence
- Parameter convergence
- Function evaluations convergence
- Constraint satisfaction
NMAopt Results

\begin{figure}
\centering
\includegraphics[width=\textwidth]{nmaopt_results.png}
\caption{Examples of running constrained optimization codes.}
\end{figure}
SQPopt Results

The SQPopt results show the optimization process for the objective function $f_{\text{aug}}(x_1, x_2)$ subject to constraints. The optimal solution is found at $f_{\text{opt}} = 1.0151 \times 10^0$ with maximum constraint violation $\max(g_{\text{opt}}) = 2.3184 \times 10^{-4}$. The convergence plots indicate the progression of the objective function $F$, the parameter values, and the maximum constraints with respect to function evaluations.
Optimization of Uncertain Systems

The design-based analysis of many systems usually involves imprecise or uncertain information. These can include uncertainties in the system’s operating environment (loading such as the monthly rainfall into a watershed, the peak hurricane wind loading on a skyscraper, or the impact velocity of a car crash), and uncertainties in the system’s intrinsic attributes (behavior such as watershed dynamics or material strength).

Probability distributions (e.g., the mean and standard deviation of a normal distribution), provide a quantification of these intrinsic and environmental uncertainties. A distribution of a system’s performance can be estimated by computing a statistical sample of the performance from samples of intrinsic and environmental variables, e.g., a sample of material strength values, and sample of peak wind speed values. This is done through the process of multiple re-analyses of a candidate design using different, but statistically representative, values of the uncertain quantities in each analysis. The mean and variability of the performance of a particular design (in the context of intrinsic and environmental uncertainties) can be estimated from a statistical sample of the performance metric (i.e., the objective function). Note that confidence in the statistical estimates improves with the number of re-analyses. In other words, estimates of statistics like the mean or the standard deviation improve with larger sample sizes. System analyses can take considerable computational time, and so it is desirable to limit the number of re-analyses, even at the cost of poorer estimates of the mean and standard deviation of the performance.

Now, comparing two candidate designs in terms of poor estimates of the mean and the standard deviation of the performance, could be misleading. The comparative assessment based on statistics estimated from a small sample of performance metrics depends entirely on the values in the sample, which might not be truly statistically representative if the sample size is small. Design A could appear to be better than design B based on a small sample, while a very large sample would indicate the opposite.

To optimize designs with uncertain performance, search methods (such as the Nelder-Mead method) can be more robust (less sensitive to sampling variability) than gradient-based methods (such as SQP).

The implementation of the Optimized Step Size Random Search in ORSopt and the Nelder-Mead method in NMAopt.m can handle problems with uncertain performance metrics. The use of these optimization methods for optimization with uncertainty can be tailored (or tuned) by changing three values in the set of options:

- Uncertain objective functions can be assessed in terms of their statistical properties (for example, their mean (average) and their standard deviation (or coefficient of variation)). Estimating the mean performance \( M_F(x) \) and coefficient of variation of the performance \( C_F(x) \) requires a sample of values of the performance metric \( f \) for the same set of design variables \( x \). Given a sample \([f_1, f_2, ..., f_m]\) from \( m \) repeated evaluations of the objective function, the estimates of the mean and coefficient of variation can be computed as follows:

\[
M_F = \frac{1}{m} \sum_{i=1}^{m} f_i
\]

and

\[
C_F = \frac{1}{M_F} \left[ \frac{1}{m-1} \sum_{i=1}^{m} (f_i - M_F)^2 \right]^{1/2}
\]

Because the estimate of the mean and c.o.v. of \( F \) requires a sample of \( m \) evaluations, the total number of evaluations for the optimization will be larger than what is used in deterministic optimization (without uncertainty). To allow for this the maximum total number of function evaluations, MaxEvals (the fifth value in options) should be set larger.

- How many evaluations should be used in computing \( M_F \) and \( C_F \)? That is, what value for \( m \) should be specified? The answer to this question depends on the level of inherent variability in \( f(x) \), on the desired statistical error \( \epsilon_F \) in the estimate of \( M_F \), and the level of confidence we require of our estimate. The inherent variability of \( f(x) \) is estimated as \( C_F \).
The greater this inherent variability, the larger the sample size should be.
- The smaller the desired error, the larger the sample size should be.
- The larger the level of confidence in the estimate, the larger the sample size should be.

The equation for $m$ is:

$$m = \left[ \frac{z_{\alpha/2}}{e_F} \right]^2$$

where $z_{\alpha/2} = 1.645$ for a 90% confidence level on the estimate of $M_F$.

The ninth value in options, mMax, sets a limit on the sample size $m$ in order to restrict $m$ from becoming too large, (in cases with very large $C_F$ or very small desired $e_F$).

The tenth value in options, errF, sets the value of $e_F$, the desired estimation error for the mean, $M_F$.

Note that optimizing with small values of errF and large values of mMax could require many (many) function evaluations, but will represent the statistics of the objective function very well.

- For optimization problems with uncertain objective functions, it is sometimes desirable to recognize the uncertainty of the objective in the cost function. Optimization cost functions for uncertain objective functions are called risk measures.

The .m-file avg_cov_func.m takes care of computing $M_F$, $C_F$, and the risk measure to be optimized. A number of risk measures for stochastic optimization problems can be selected within avg_cov_func.m: the sample max, the sample average, the sample average plus the sample standard deviation (the 84th percentile of the objective function), or the sample average plus the sample standard deviation divided by $\sqrt{m}$ (the 84th percentile of the mean estimate).

In the example script below, the uncertainty-level in this example problem is set by the coefficient $c_3$, so in this example we know in advance that the standard deviation of $F$ is equal to the value we use for $c_3$ (0.10). Setting errF to 0.05 means that we desire an estimate for the mean of $F$ that is accurate to within $\pm 5\%$, with a 90% confidence level. Using this information along with the equation for $m$, above, we will need a sample size of $m = (1.645 \times 0.10/0.05)^2 \approx 11$. We have set the maximum sample size, mMax, equal to 11.

Overall, the goal in setting values of mMax and errF is to use values that get the overall optimization to consistently converge to good solutions with the smallest number of function evaluations.

Note that:

- The risk measure used in this example is $F_{\text{risk}} = M_F(1 + C_F)$.
- The values of the optimized objective functions $f_{\text{opt}}$ shown in the figures are all very close to one another, even for the problem with added uncertainty.
- The uncertain objective function takes about eleven times as many function evaluations to converge as does the problem without uncertainty. This is related to the value of mMax which is set to 11 in the uncertain case and is set to 1 in the case without uncertainty.
% optim_example_optU.m — Optimization with Uncertainty

% H.P. Gavin, Dept. Civil & Environ. Eng’g, Duke Univ., 2011–12

x_lb = [ 0 ; 0 ];   % lower bound on permissible parameter values
x_ub = [ 1 ; 1 ];   % upper bound on permissible parameter values

x_init = x_lb + rand(2,1).*(x_ub - x_lb); % random initial guess, or ...
x_init = [ 0.8 ; 0.8 ]; % a specified initial guess

cstnts = [ 0.8 ; 0.2 ; 0.10 ];  % other constants used in the cost function

% algorithmic parameters ...
% display tolX tolF tolG MaxEvals Penalty Exponent mMax errF
options = [ 4 0.05 0.05 0.001 5e3 1.0 0.5 11 0.05 ];
% ... ORSopt or NMAopt or SQPopt ...
[x_opt, f_opt, g_opt, cvg_hst] = NMAopt('optim_example_analysis', x_init, x_lb, x_ub, options, cstnts);

plot_cvg_hst( cvg_hst, x_opt, 100); % plot the convergence history
Figure 2. NMAopt convergence on penalized objective function surface with 15% uncertainty in \( f(x) \).

\[
\begin{align*}
\text{Objective: } f_{\text{opt}} &= 1.0102e+00 \\
\max(g_{\text{opt}}) &= -7.0405e-02
\end{align*}
\]

Figure 3. NMAopt convergence histories.
Examples of Running Constrained Optimization Codes

ORSopt.m

```matlab
% [x_opt, f_opt, g_opt, cvg_hst] = ORSopt(func, x_init, x_lb, x_ub, options, consts)

% ORSopt: Optimized Step Size Randomized Search
% Nonlinear optimization with inequality constraints using Random Search
% minimizes f(x) such that g(x) <= 0 and x_lb <= x_opt <= x_ub.
% f is a scalar objective function, x is a vector of design variables, and
% g is a vector of constraints.

% INPUT
% func : the name of the matlab function to be optimized in the form
% [objective, constraints] = func(x, consts)
% x_init : the vector of initial parameter values
% x_lb : lower bound on permissible values of the variables, x (n x 1)
% x_ub : upper bound on permissible values of the variables, x (n x 1)
% options : options(1) = 1 means display intermediate results
% options(2) = tol_x tolerance on convergence of design variables
% options(3) = tol_f tolerance on convergence of objective
% options(4) = tol_g tolerance on convergence of constraints
% options(5) = max_eval_limit on number of function evaluations
% options(6) = penalty on constraint violations
% options(7) = exponent on constraint violations
% options(8) = max number of function evals in est. of mean f(x)
% options(9) = desired accuracy of mean f (as a c.o.v.)
% options(10) = 1 means stop when solution is feasible
% consts : an optional vector of constants to be passed to func(x, consts)

% OUTPUT
% x_opt : a set of design variables at or near the optimal value
% f_opt : the objective associated with the near-optimal design variables
% g_opt : the constraints associated with the near-optimal design variables
% cvg_hst : record of x, f, g, function_count, and convergence criteria
```

NMAopt.m

```matlab
% [x_opt, f_opt, g_opt, cvg_hst] = NMAopt(func, x_init, x_lb, x_ub, options, consts)

% NMAopt: Nelder-Mead method for the nonlinear optimization with inequality constraints
% minimizes f(x) such that g(x) <= 0 and x_lb <= x_opt <= x_ub.
% f is a scalar objective function, x is a vector of design variables, and
% g is a vector of constraints.

% INPUT
% func : the name of the matlab function to be optimized in the form
% [objective, constraints] = func(x, consts)
% x_init : the vector of initial parameter values
% x_lb : lower bound on permissible values of the variables, x (n x 1)
% x_ub : upper bound on permissible values of the variables, x (n x 1)
% options : options(1) = 1 means display intermediate results
% options(2) = tol_x tolerance on convergence of variables
% options(3) = tol_f tolerance on convergence of objective
% options(4) = tol_g tolerance on convergence of constraints
% options(5) = max_eval_limit on number of function evaluations
% options(6) = penalty on constraint violations
% options(7) = exponent on constraint violations
% options(8) = max number of function evals in est. of mean f(x)
% options(9) = desired accuracy of mean f (as a c.o.v.)
% options(10) = 1 means stop when solution is feasible
% consts : an optional vector of values that are not design variables

% OUTPUT
% x_opt : a set of design variables at or near the optimal value
% f_opt : the objective associated with the near-optimal design variables
% g_opt : the constraints associated with the near-optimal design variables
% cvg_hst : record of x, f, g, function_count, and convergence criteria
```
SQPopt.m

```matlab
% [x_opt, f_opt, g_opt, cvg_hst, lambda, Hess]=SQPopt(func, x_init, x_lb, x_ub, options, c)
% SQPopt: Nonlinear optimization with inequality constraints using Sequential Quadratic Programming
% minimizes f(x) such that g(x)<0 and x_lb <= x_opt <= x_ub.
% f is a scalar objective function, p is a vector of design variables,
% and g is a vector of constraints.
% INPUT
% = = = = = =
% func : the name of the matlab function to be optimized in the form [objective, constraints] = func(x, c)
% x_init : the vector of initial parameter values (n x 1)
% x_lb : lower bound on permissible values of the variables, x (n x 1)
% x_ub : upper bound on permissible values of the variables, x (n x 1)
% options : options(1) = 1 means display intermediate results
% options(2) = tol_p tolerance on convergence of variables
% options(3) = tol_f tolerance on convergence of objective
% options(4) = tol_g tolerance on constraint functions
% options(5) = max_evals limit on number of function evaluations
% c : an optional vector of constants used by func(x, c)
% OUTPUT
% = = = = = =
% x_opt : a set of design variables at or near the optimal value
% f_opt : the objective associated with the optimal design variables
% g_opt : the constraints associated with the optimal design variables
% lambda : the set of Lagrange multipliers at the active constraints
% Hess : the Hessian of the objective function at the optimal point
```