Methods for Constrained Optimization

Dr. José Ernesto Rayas-Sánchez

1

Outline

- Constrained optimization problem
- Box constraints
- Methods for constrained optimization problems
- Elimination of variables
- Penalty methods
- Sequential quadratic programming (SQP)
- Minimax formulations

Dr. J. E. Rayas-Sánchez

2

Constrained Optimization Problem

Standard form:

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} u(\boldsymbol{x})$$

subject to

$$\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$$

$$\boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}$$

$$\boldsymbol{x}^{\text{lb}} \leq \boldsymbol{x} \leq \boldsymbol{x}^{\text{ub}}$$

- $\boldsymbol{x}, \boldsymbol{x}^{\text{lb}}, \boldsymbol{x}^{\text{ub}} \in \Re^n$
- $u: \Re^n \rightarrow \Re$, $\boldsymbol{h}: \Re^n \rightarrow \Re^E$, $\boldsymbol{g}: \Re^n \rightarrow \Re^I$
- It is assumed $n > E$

Constrained Optimization Problem (cont)

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} u(\boldsymbol{x})$$

subject to

$$\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$$

$$\boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}$$

$$\boldsymbol{x}^{\text{lb}} \leq \boldsymbol{x} \leq \boldsymbol{x}^{\text{ub}}$$

It is generally assumed that satisfying all the constraints is more important than minimizing $u(\boldsymbol{x})$, i.e., feasibility is more important than optimality

Constrained Optimization Problem (cont)

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} u(\boldsymbol{x})$$

subject to

$$\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$$

$$\boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}$$

$$\boldsymbol{x}^{\text{lb}} \leq \boldsymbol{x} \leq \boldsymbol{x}^{\text{ub}}$$

The feasible set:

$$\Omega = \{\boldsymbol{x} \in \Re^n \mid \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \wedge \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0} \wedge \boldsymbol{x}^{\text{lb}} \leq \boldsymbol{x} \leq \boldsymbol{x}^{\text{ub}}\}$$

Constrained Optimization Problem (cont)

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}\in\Omega} u(\boldsymbol{x})$$

$$\Omega = \{\boldsymbol{x} \in \Re^n \mid \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \wedge \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0} \wedge \boldsymbol{x}^{\text{lb}} \leq \boldsymbol{x} \leq \boldsymbol{x}^{\text{ub}}\}$$

Dealing with Box Constraints

- Box constraints can be treated as inequality constraints

$$\mathbf{x}^{\text{lb}} \le \mathbf{x} \le \mathbf{x}^{\text{ub}} \quad \rightarrow \quad \begin{aligned} g_1(\mathbf{x}) &= x_1 - x_1^{\text{ub}} \le 0 \\ g_2(\mathbf{x}) &= x_1^{\text{lb}} - x_1 \le 0 \\ &\vdots \\ g_{2n-1}(\mathbf{x}) &= x_n - x_n^{\text{ub}} \le 0 \\ g_{2n}(\mathbf{x}) &= x_n^{\text{lb}} - x_n \le 0 \end{aligned}$$

- They can also be considered by restricting the optimization space (through variable transformations)

Dr. J. E. Rayas-Sánchez 7

Box Constraints – Restricting Optimization Space

- Box constraints can be incorporated into an unconstrained optimization problem by transforming the optimization variables
- Instead of solving

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} u(\mathbf{x})$$

subject to

$$\mathbf{x}^{\text{lb}} \le \mathbf{x} \le \mathbf{x}^{\text{ub}}$$

we solve

$$\mathbf{z}^* = \arg\min_{\mathbf{z}} u(\mathbf{z})$$

Constraint	Transformation
$x_i \ge 0$	$x_i = z_i^2$
$x_i > 0$	$x_i = e^{z_i}$
$x_i \ge x_i^{\text{lb}}$	$x_i = x_i^{\text{lb}} + z_i^2$
$x_i > x_i^{\text{lb}}$	$x_i = x_i^{\text{lb}} + e^{z_i}$
$-1 \le x_i \le 1$	$x_i = \sin z_i$
$0 \le x_i \le 1$	$x_i = (\sin z_i)^2$
$0 < x_i < 1$	$x_i = \dfrac{e^{z_i}}{1 + e^{z_i}}$

(Bandler, 1997)

Dr. J. E. Rayas-Sánchez 8

Box Constraints – Restricting Opt. Space (cont)

Constraint	Transformation
$x_i^{\text{lb}} \le x_i \le x_i^{\text{ub}}$	$x_i = x_i^{\text{lb}} + (x_i^{\text{ub}} - x_i^{\text{lb}})(\sin z_i)^2$
	$x_i = \frac{1}{2}\left(x_i^{\text{lb}} + x_i^{\text{ub}}\right) + \frac{1}{2}(x_i^{\text{ub}} - x_i^{\text{lb}})\sin z_i$
$x_i^{\text{lb}} < x_i < x_i^{\text{ub}}$	$x_i = x_i^{\text{lb}} + \left(x_i^{\text{ub}} - x_i^{\text{lb}}\right)\left(\frac{e^{z_i}}{1+e^{z_i}}\right)$

(Bandler, 1997)

Methods for Constrained Optimization

- Indirect methods (or Sequential Unconstrained Minimization Techniques, SUMT):
  - Elimination of variables (equality constraints)
  - Exterior penalty function method (EPF)
  - Augmented Lagrange multiplier method (ALM)
- Direct methods:
  - Sequential linear programming (SLP)
  - Sequential quadratic programming (SQP)
  - Generalized reduced gradient method (GRG)
  - Sequential gradient restoration algorithm (SGRA)

Equality Constraints – Elimination of Variables

- When solving

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} u(\mathbf{x})$$

subject to

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}$$

  we can reduce the number of equality constraints by eliminating some of the optimization variables

- If sufficient variables are eliminated, we can obtain an unconstrained optimization problem
- This technique must be carefully used (the resultant problem can be ill-conditioned)

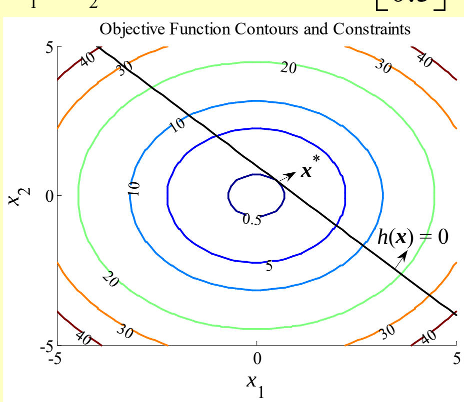
Elimination of Variables – Example 1 ☺

$$\min_{\mathbf{x}} x_1^2 + x_2^2$$

subject to

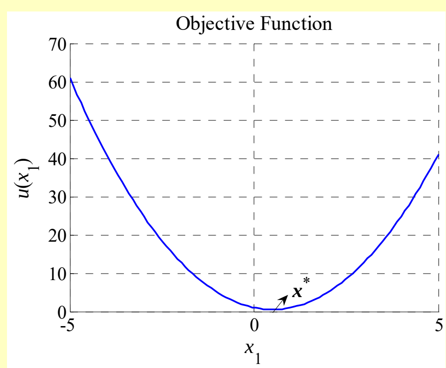
$$x_1 + x_2 - 1 = 0$$

$$\mathbf{x}^* = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$



→

$$\min_{\mathbf{x}} x_1^2 + (1 - x_1)^2$$



$$x_2^* = 1 - x_1^*$$

Elimination of Variables – Example 2 ☹

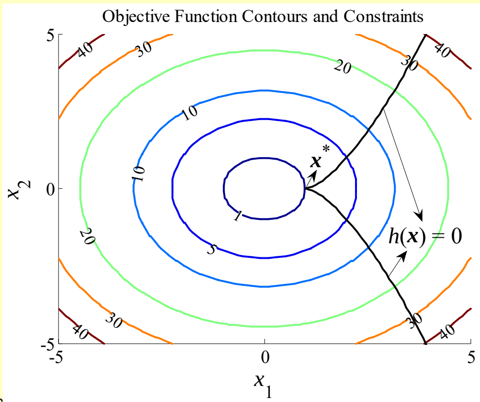
$$\min_{\boldsymbol{x}} x_1^2 + x_2^2$$

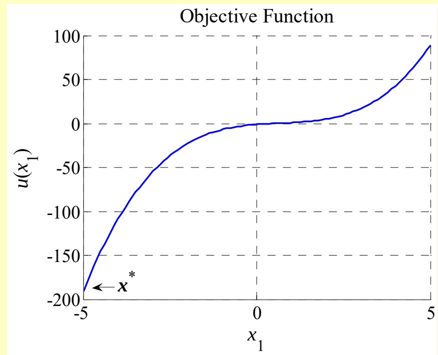
subject to

$$(x_1 - 1)^3 - x_2^2 = 0$$

$$\boldsymbol{x}^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\min_{\boldsymbol{x}} x_1^2 + (x_1 - 1)^3$$





Minimizer is unbounded

Equality Constraints – Penalty Functions

Instead of solving

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} u(\boldsymbol{x})$$

subject to

$$\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$$

we solve

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} U(\boldsymbol{x})$$

where

$$U(\boldsymbol{x}) = u(\boldsymbol{x}) + r^{\text{h}} \|\boldsymbol{h}(\boldsymbol{x})\|_2^2$$

$\|\boldsymbol{h}(\boldsymbol{x})\|_2^2$: penalty function

$r^{\text{h}} \in \Re$: penalty coefficient

Inequality Constraints – Penalty Functions

Instead of solving

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} u(\boldsymbol{x})$$

subject to

$$\boldsymbol{g}(\boldsymbol{x}) \le \boldsymbol{0}$$

we solve

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} U(\boldsymbol{x})$$

where

$$U(\boldsymbol{x}) = u(\boldsymbol{x}) + r^{\mathrm{g}} \|\boldsymbol{G}(\boldsymbol{x})\|_2^2$$

$$G_j = \max\{0, g_j(\boldsymbol{x})\}$$

$\|\boldsymbol{G}(\boldsymbol{x})\|_2^2$: penalty function

$r^{\mathrm{g}} \in \Re$: penalty coefficient

Exterior Penalty Function (EPF) Method

Original problem:

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} u(\boldsymbol{x})$$

subject to

$$\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$$

$$\boldsymbol{g}(\boldsymbol{x}) \le \boldsymbol{0}$$

$$\boldsymbol{x}^{\mathrm{lb}} \le \boldsymbol{x} \le \boldsymbol{x}^{\mathrm{ub}}$$

Indirect solution through EPF method:

$$\boldsymbol{z}^* = \arg\min_{\boldsymbol{z}} u(\boldsymbol{z}) + p(\boldsymbol{z}, r^{\mathrm{h}}, r^{\mathrm{g}})$$

$$p(\boldsymbol{z}, r^{\mathrm{h}}, r^{\mathrm{g}}) = r^{\mathrm{h}} \|\boldsymbol{h}(\boldsymbol{z})\|_2^2 + r^{\mathrm{g}} \|\boldsymbol{G}(\boldsymbol{z})\|_2^2$$

$$G_j = \max\{0, g_j(\boldsymbol{z})\}$$

$$x_i = x_i^{\mathrm{lb}} + (x_i^{\mathrm{ub}} - x_i^{\mathrm{lb}})(\sin z_i)^2$$

$$j = 1,2,\ldots,I \qquad i = 1,2,\ldots,n$$

Exterior Penalty Function (EPF) Method (cont)

- The optimal solution $\boldsymbol{z}^* = \arg\min_{\boldsymbol{z}} u(\boldsymbol{z}) + p(\boldsymbol{z}, r^{\text{h}}, r^{\text{g}})$

  is a function of the penalty coefficients r^{h} and r^{g}
- Penalty coefficients should be gradually increased until all constraints are satisfied (exterior method)
- The EPF method is very sensitive to the initial values of the penalty coefficients, r^{h} and r^{g}

An EPF Algorithm

$\boldsymbol{x}^* = \text{EPF}(u, \boldsymbol{x}_0, \boldsymbol{h}, \boldsymbol{g})$

$u: \Re^n \rightarrow \Re$; $\boldsymbol{h}: \Re^n \rightarrow \Re^E$; $\boldsymbol{g}: \Re^n \rightarrow \Re^I$;
$\boldsymbol{x}_0, \boldsymbol{x}^* \in \Re^n$

begin
  set c^{h}, c^{g}, and ε

  $r_0^{\text{h}} = \dfrac{|u(\boldsymbol{x}_0)|}{\|\boldsymbol{h}(\boldsymbol{x}_0)\|_2^2 + \varepsilon}$; $r_0^{\text{g}} = \dfrac{|u(\boldsymbol{x}_0)|}{\|\boldsymbol{g}(\boldsymbol{x}_0)\|_2^2 + \varepsilon}$; $i = 0$

  repeat until *StoppingCriteria*

    $\boldsymbol{x}_0 = \boldsymbol{x}_i$

    $\boldsymbol{x}_{i+1} = \arg\min_{\boldsymbol{x}} u(\boldsymbol{x}) + r_i^{\text{h}} \sum_{k=1}^{E} h_k^2(\boldsymbol{x}) + r_i^{\text{g}} \sum_{k=1}^{I} (\max\{0, g_k(\boldsymbol{x})\})^2$

    $r_{i+1}^{\text{h}} = c^{\text{h}} r_i^{\text{h}}$

    $r_{i+1}^{\text{g}} = c^{\text{g}} r_i^{\text{g}}$

    $i = i + 1$
  end
  $\boldsymbol{x}^* = \boldsymbol{x}_i$
end

Sequential Quadratic Programming (SQP)

- SQP are considered the state-of-the-art in nonlinear programming
- At each iteration, the objective function is approximated by a quadratic function, and the nonlinear constraints are approximated by linear constraints
- The quadratic sub-problem is solved to find a search direction at the current iterate
- The next iterate is obtained from a line search

SQP Sub-problem

At the current iterate $\boldsymbol{x}_i$,

- The objective function is expanded quadratically

$$u^{(i)}(\boldsymbol{d}) = u(\boldsymbol{x}_i) + \boldsymbol{d}^T \nabla u(\boldsymbol{x}_i) + \frac{1}{2}\boldsymbol{d}^T \boldsymbol{H}(u(\boldsymbol{x}_i))\boldsymbol{d} \approx u(\boldsymbol{x}_i + \boldsymbol{d})$$

- The constraints are expanded linearly

$$\boldsymbol{h}^{(i)}(\boldsymbol{x}) = \boldsymbol{h}(\boldsymbol{x}_i) + \boldsymbol{J}(\boldsymbol{h}(\boldsymbol{x}_i))\boldsymbol{d} \approx \boldsymbol{h}(\boldsymbol{x}_i + \boldsymbol{d})$$

$$\boldsymbol{g}^{(i)}(\boldsymbol{x}) = \boldsymbol{g}(\boldsymbol{x}_i) + \boldsymbol{J}(\boldsymbol{g}(\boldsymbol{x}_i))\boldsymbol{d} \approx \boldsymbol{g}(\boldsymbol{x}_i + \boldsymbol{d})$$

SQP Sub-problem (cont)

At the current iterate $\boldsymbol{x}_i$,

- The search direction $\boldsymbol{d}_i$ is found by solving

$$\boldsymbol{d}_i = \arg\min_{\boldsymbol{d}} \boldsymbol{d}^T \nabla u(\boldsymbol{x}_i) + \frac{1}{2}\boldsymbol{d}^T \boldsymbol{H}(u(\boldsymbol{x}_i))\boldsymbol{d}$$

subject to

$$\boldsymbol{h}(\boldsymbol{x}_i) + \boldsymbol{J}(\boldsymbol{h}(\boldsymbol{x}_i))\boldsymbol{d} = \boldsymbol{0}$$
$$\boldsymbol{g}(\boldsymbol{x}_i) + \boldsymbol{J}(\boldsymbol{g}(\boldsymbol{x}_i))\boldsymbol{d} \le \boldsymbol{0}$$
$$\boldsymbol{x}^{\text{lb}} \le \boldsymbol{x} \le \boldsymbol{x}^{\text{ub}}$$

- The next iterate is calculated using $\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \alpha^*\boldsymbol{d}_i$ where α^* is obtained from a line search on $u(\boldsymbol{x}_i + \alpha\boldsymbol{d}_i)$

Dr. J. E. Rayas-Sánchez 21

Minimax Formulations

- Minimax formulations are used to minimize the maximum error of a function (model response) with respect to a number of specifications
- A minimax formulation can be implemented as a constrained or as an unconstrained optimization problem

Dr. J. E. Rayas-Sánchez 22

Minimax Formulations – Unconstrained

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} \max\{\ldots e_k(\boldsymbol{x})\ldots\}$$

where

$$e_k(\boldsymbol{x}) = \begin{cases} R_k(\boldsymbol{x}) - S_k^{\text{ub}} & \text{for all } k \in I^{\text{ub}} \\ S_k^{\text{lb}} - R_k(\boldsymbol{x}) & \text{for all } k \in I^{\text{lb}} \end{cases}$$

- $R_k(\boldsymbol{x})$ is the k-th model response at point $\boldsymbol{x}$
- S_k^{ub} and S_k^{lb} are the upper and lower bound specifications
- I^{ub} and I^{lb} are index sets (not necessarily disjoint)
- Vector $\boldsymbol{e}(\boldsymbol{x})$ contains all the error functions with respect to the design specifications

Dr. J. E. Rayas-Sánchez 23

Minimax Formulations – Unconstrained (cont)

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} U(\boldsymbol{x})$$

$$U(\boldsymbol{x}) = \max\{\ldots e_k(\boldsymbol{x})\ldots\}$$

where

$$e_k(\boldsymbol{x}) = \begin{cases} R_k(\boldsymbol{x}) - S_k^{\text{ub}} & \text{for all } k \in I^{\text{ub}} \\ S_k^{\text{lb}} - R_k(\boldsymbol{x}) & \text{for all } k \in I^{\text{lb}} \end{cases}$$

Dr. J. E. Rayas-Sánchez 24

Minimax Formulations – Unconstrained (cont)

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} U(\boldsymbol{x})$$

$$U(\boldsymbol{x}) = \max\{\ldots e_k(\boldsymbol{x})\ldots\}$$

Equality specifications S_k^{eq} can be implemented as a combination of upper and lower specifications ($\Delta S^{\text{eq}} > 0$)

$$e_k(\boldsymbol{x}) = \begin{cases} R_k(\boldsymbol{x}) - (S_k^{\text{eq}} + \Delta S^{\text{eq}}) & \text{for all } k \in I^{\text{eq}} \\ (S_k^{\text{eq}} - \Delta S^{\text{eq}}) - R_k(\boldsymbol{x}) & \text{for all } k \in I^{\text{eq}} \end{cases}$$

or as a single error function

$$e_k(\boldsymbol{x}) = \left| R_k(\boldsymbol{x}) - S_k^{\text{eq}} \right| - \Delta S^{\text{eq}} \quad \text{for all } k \in I^{\text{eq}}$$

Dr. J. E. Rayas-Sánchez 25

Minimax Formulations – Unconstrained (cont)

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} \max\{\ldots e_k(\boldsymbol{x})\ldots\}$$

where

$$e_k(\boldsymbol{x}) = \begin{cases} R_k(\boldsymbol{x}) - S_k^{\text{ub}} & \text{for all } k \in I^{\text{ub}} \\ S_k^{\text{lb}} - R_k(\boldsymbol{x}) & \text{for all } k \in I^{\text{lb}} \\ \left| R_k(\boldsymbol{x}) - S_k^{\text{eq}} \right| - \Delta S^{\text{eq}} & \text{for all } k \in I^{\text{eq}} \end{cases}$$

- $R_k(\boldsymbol{x})$ is the k-th model response at point $\boldsymbol{x}$
- S_k^{ub} and S_k^{lb} are the upper and lower bound specifications
- S_k^{eq} is an equality specification ($\pm\Delta S^{\text{eq}}$)
- I^{ub} and I^{lb} are index sets (not necessarily disjoint)

Dr. J. E. Rayas-Sánchez 26

Minimax Formulations – Relative Errors

- Formulation

$$e_k(\boldsymbol{x}) = \begin{cases} R_k(\boldsymbol{x}) - S_k^{\text{ub}} & \text{for all } k \in I^{\text{ub}} \\ S_k^{\text{lb}} - R_k(\boldsymbol{x}) & \text{for all } k \in I^{\text{lb}} \\ \left| R_k(\boldsymbol{x}) - S_k^{\text{eq}} \right| - \Delta S^{\text{eq}} & \text{for all } k \in I^{\text{eq}} \end{cases}$$

may require the usage of weighting factors

- We can use instead a relative formulation for the error functions

Dr. J. E. Rayas-Sánchez 27

Minimax Formulations – Relative Errors (cont)

- Using relative error functions (assuming $S_k^{\text{ub}} > 0$ and $S_k^{\text{lb}} > 0$)

$$e_k(\boldsymbol{x}) = \begin{cases} \dfrac{R_k(\boldsymbol{x})}{S_k^{\text{ub}} + \varepsilon} - 1 & \text{for all } k \in I^{\text{ub}} \\ 1 - \dfrac{R_k(\boldsymbol{x})}{S_k^{\text{lb}} + \varepsilon} & \text{for all } k \in I^{\text{lb}} \\ \dfrac{| R_k(\boldsymbol{x}) - S_k^{\text{eq}} |}{\varepsilon} - 1 & \text{for all } k \in I^{\text{eq}} \end{cases}$$

where ε is an arbitrary small positive number

Dr. J. E. Rayas-Sánchez 28

Minimax Formulations – Constrained

- We define an additional optimization variable (ceiling)

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} x_{n+1}$$

subject to

$$e_k(\boldsymbol{x}) - x_{n+1} \leq 0$$

where
$$e_k(\boldsymbol{x}) = \begin{cases} R_k(\boldsymbol{x}) - S_k^{\text{ub}} & \text{for all } k \in I^{\text{ub}} \\ S_k^{\text{lb}} - R_k(\boldsymbol{x}) & \text{for all } k \in I^{\text{lb}} \end{cases}$$

- $R_k(\boldsymbol{x})$ is the k-th model response at point $\boldsymbol{x}$
- S_k^{ub} and S_k^{lb} are the upper and lower bound specifications
- I^{ub} and I^{lb} are index sets (not necessarily disjoint)