第一章 非线性规划

1.1 非线性规划问题的数学模型

记 $x=(x_1,x_2,\ldots,x_n)^T\in R^n$，$f(x)$、$g_i(x)$ 和 $h_j(x)$ 为定义在 R^n 上的实值函数。若 $f(x)$，$g_i(x)$, $i=1,2,\ldots,n$ 和 $h_j(x), j=1,2,\ldots,q$ 中至少有一个为 x 的非线性函数，称如下的模型：

$$\min\ f(x)$$
$$\text{s.t.}\begin{cases} g_i(x)\le 0,\ i=1,2,\ldots,p, \\ h_j(x)=0,\ j=1,2,\ldots,q \end{cases}$$

为非线性规划模型。

该模型也可表示为向量形式：

$$\min\ f(x),$$
$$\text{s.t.}\begin{cases} G(x)\le 0 \\ H(x)=0 \end{cases}$$

其中，$G(x)=(g_1(x),g_2(x),\ldots,g_p(x))^T$，$H(x)=(h_1(x),h_2(x),\ldots,h_q(x))^T$。

定义 1.1 (可行域)

满足所有约束条件的点 $x\in\mathbb{R}$ 的集合，记为 K。

♣

定义 1.2 (全局最优解)

若 $x^*\in K$，且对任意的 $x\in K\backslash\{x^*\}$，都有：

$$f(x^*)\le f(x)$$

成立，则称 $f(x^*)$ 为其全局最优值。

♣

若不等式严格成立，则称 x^* 为严格全局最优解，$f(x^*)$ 为严格全局最优值。

定义 1.3 (局部最优解)

若 $x^*\in K$，且对任意的 $x\in N_\delta\cap K$，都有：

$$f(x^*)\le f(x)$$

成立，称 $f(x^*)$ 为其全局最优值。

♣

若不等式严格成立，则称 x^* 为严格局部最优解，$f(x^*)$ 为严格局部最优值。

1.2 无约束非线性规划

1.2.1 数学模型

无约束非线性规划问题可表示为：

$$\min\ f(x),\ x \in R^n$$

求二元函数极值的方法可以推广到无约束优化问题中。

命题 1.1

设 $f(x)$ 具有连续的一阶偏导数。若 $x = x^*$ 是无约束问题的局部极小点，则 $f(x)$ 的梯度 $\nabla f(x)$ 满足：

$$\nabla f(x^*) = 0.$$

命题 1.2

设 $f(x)$ 具有连续的二阶偏导数，在 $x = x^*$ 处满足：

$$\nabla f(x^*) = 0,$$

且 Hessian 矩阵 $\nabla^2 f(x^*)$ 正定，则 $f(x^*)$ 为无约束优化问题的局部最优解，其中，

$$\nabla^2 f(x^*) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

说明：求解无约束优化问题困难的是求解方程 $\nabla f(x^*) = 0$。对于复杂的函数，常用数值解法，如最速下降法、牛顿法和拟牛顿法等。

实际应用的绝大多数优化问题都是有约束的。非线性规划目前还没有一般性的算法，各个算法都有其特定的适用范围，且都有一定的局限性。

1.3 Matlab 求解无约束非线性规划问题

用于求解无约束极小值问题的函数有：fminsearch 和 fminunc。

例题 1.1 求函数 $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$ 的极小点。

调用 fminsearch 命令的 MATLAB 程序如下。

```
clc
clear
f = @(x) 100*(x(2)-x(1)^2)^2 + (1-x(1))^2;
x0 = [0, 0];
[x, fval] = fminsearch(f, x0)
```

计算结果：

$$x = [1, 1],\ \text{fval} = 3.6862e - 10.$$

调用 fminunc 命令的 MATLAB 程序如下。

```
clc
clear
f = @(x) 100*(x(2) -x(1)^2)^2 + (1-x(1))^2;
x0 = [0, 0];
[x, fval] = fminunc(f, x0)
```

计算结果：

$$x = [1, 1],\ \text{fval} = 1.9474e - 11.$$

1.4 有约束非线性规划

总体思路：尽可能将非线性问题转化为线性问题，将有约束问题转化为无约束问题。对于特殊的只有等式约束的非线性规划问题的情形：

$$\min\ f(x),$$
$$\text{s.t.} \begin{cases} h_j(x) = 0,\ j = 1, 2, \ldots, q \\ x \in R^n \end{cases}$$

结论（Lagrange 定理）设函数 $f(x), h_1(x), h_2(x), \ldots, h_q(x)$ 在可行点 x^* 的某个邻域内可微，向量组 $\nabla h_j(x^*)$ 线性无关，令

$$L(x, \lambda) = f(x) + \lambda^T H(x)$$

其中，$\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_q)^T$，$H(x) = (h_1(x), h_2(x), \ldots, h_q(x))^T$。若 x^* 为局部最优解，则存在实向量 $\lambda^* = (\lambda_1^*, \lambda_2^*, \ldots, \lambda_q^*)^T$，使得：

$$\nabla L(x^*, \lambda^*) = 0$$

即：

$$\nabla f(\mathbf{x}^*) + \sum_{j=1}^{q} \lambda_j^* \nabla h_j(\mathbf{x}^*) = 0$$

.

对于存在不等式约束的非线性规划问题，无法直接应用 Lagrange 定理将其转化为无约束问题。利用罚函数可以实现问题的无约束转化，但该方法的计算精度可能较差，一般用软件工具库求解非线性规划问题。

1.5 有约束非线性规划问题的求解

有约束非线性规划问题的 Matlab 求解方法有两种：基于求解器的求解方法和基于问题的求解方法。

$\boxed{\text{基于求解器的解法}}$

将非线性规划的数学模型改写为标准形式：

$$\min\ f(x),$$
$$\text{s.t.}\begin{cases} A\cdot x\le b\\ \text{Aeq}\cdot x=\text{beq}\\ c(x)\le 0\\ \text{ceq}(x)=0\\ \text{lb}\le x\le \text{ub}\end{cases}$$

其中，f(x) 是标量函数，A, b, Aeq, beq, lb, ub 是相应维数的矩阵和向量，c(x), ceq(x) 分别是非线性不等式约束和等式约束函数。

Matlab 命令：

[x, fval] = fmincon(fun, x0, A, b, Aeq, beq, lb, ub, nonlcon)

输入的参数及含义如下表所示。

fun	用 M 函数或匿名函数定义的函数
x0	x 的初始值
A, b	线性不等式约束：Ax≤ b；若无，则：A = [ ], b = [ ]
Aeq, beq	线性等式约束：Aeq x = beq；若无，则：Aeq = [ ], beq = [ ]
lb	决策向量 x 的下界，若无，则：lb = [ ]，或 lb = -inf
ub	决策向量 x 的上界，若无，则：ub = [ ]，或 ub = inf
nonlcon	用 M 函数定义的非线性向量函数

输出：x 为决策向量的取值，fval 目标函数的取值。

例题 1.2 求非线性规划：

$$f(x)=x_1^2+x_2^2+x_3^2+8$$
$$\text{s.t.}\begin{cases} x_1^2-x_2+x_3^2\ge 0\\ x_1+x_2^2+x_3^3\le 20\\ -x_1-x_2^2+2=0\\ x_2+2x_3^2=3\\ x_1,x_2,x_3\ge 0\end{cases}$$

基于求解器的程序如下。

```
clc
clear
fun1 = @(x) sum(x.^2)+8;
[x, y] = fmincon(fun1, rand(3,1), [ ], [ ], [ ], [ ], zeros(3,1), [ ], @fun2)

% 定义 fun2 函数
function [c, ceq] = fun2(x)
c=[-x(1)^2+x(2)-x(3)^2;  x(1)+x(2)^2+x(3)^3-20];
ceq=[-x(1)-x(2)^2 +2;  x(2)+2*x(3)^2-3];
end
```

求得最优解：$x = (0.5522, 1.2033, 0.9478)^T$，最小值：$f(x) = 10.6511$。
基于问题求解的程序如下。

```
clc
clear
prob = optimproblem;
x = optimvar('x', 3, 'LowerBound', 0);
prob.Objective = sum(x.^2)+8;
con1 = [-x(1)^2+x(2)-x(3)^2 <= 0, x(1)+x(2)^2+x(3)^3 <= 20];
con2 = [-x(1)-x(2)^2+2 == 0, x(2)+2* x(3)^2 == 3];
prob.Constraints.con1 = con1;
prob.Constraints.con2 = con2;
x0.x = rand(3,1);       % 非线性规划必须赋初值
[sol, fval, flag, out] = solve(prob)
sol.x           % 显示最优解
```

1.6 二次规划模型

如果规划模型的目标函数是决策向量 $x = (x_1, x_2, \ldots, x_n)^T$ 的二次函数，约束条件都是线性的，则称该模型为二次规划模型，其一般形式为：

$$\max(\min) \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_i x_j + \sum_{i=1}^{n} d_i x_i,$$

$$\text{s.t.} \begin{cases} \sum_{j=1}^{n} a_{ij}x_j \leq (\geq, =) b_i \\ x_i \geq 0,\ i = 1, 2, \ldots, n \end{cases}$$

其中, $c_{ij} = c_{ji}, i, j = 1, 2, \ldots, n$ 。

二次规划模型是一种特殊的非线性规划模型。注意到，

$$H = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

为对称矩阵。当 H 正定时，目标函数最小化时的优化问题为凸二次规划。凸二次规划的局部最优解也是全局最优解。

例题 1.3 求解如下二次规划模型：

$$\max\ -x_1^2-0.3x_1x_2-2x_2^2+98x_1+277x_2,$$
$$\text{s.t.}\begin{cases}x_1+x_2\le 100,\\ x_1-2x_2\le 0,\\ x_1,x_2\ge 0\end{cases}$$

解：将模型写为矩阵形式：

$$\max\ (x_1,x_2)\begin{pmatrix}-1 & -0.15\\ -0.15 & -2\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}+(98,277)\begin{pmatrix}x_1\\ x_2\end{pmatrix}$$
$$\text{s.t.}\begin{cases}\begin{pmatrix}1 & 1\\ 1 & -2\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}\le\begin{pmatrix}100\\ 0\end{pmatrix}\\ x_1,x_2\ge 0\end{cases}$$

该二次规划是凸规划，基于问题的求解程序如下。

```
clc
clear
prob = optimproblem('ObjectiveSense', 'max');
x = optimvar('x', 2, 'LowerBound', 0);
h = [−1, −0.15; −0.15, −2];
f=[98; 277];
prob.Objective = x' ∗ h ∗ x + f' ∗ x;
A =[1, 1; 1, -2];
b=[100; 0];
prob.Constraints = A ∗ x <= b;
[sol, fval, flag, out] = solve(prob)
sol.x           % 显示最优解
```

求得的最优解为：$x_1=35.3704, x_2=64.6296$。
目标函数的最大值为：11077.8704。

例题 1.4 求解如下二次规划模型：

$$\min\ -x_1^2 - 0.3x_1x_2 - 2x_2^2 + 98x_1 + 277x_2,$$
$$\text{s.t.}\begin{cases} x_1 + x_2 \le 100, \\ x_1 - 2x_2 \le 0, \\ x_1, x_2 \ge 0 \end{cases}$$

解：上述二次规划的目标数不是凸函数，基于问题的求解程序如下。

```
clc
clear
prob = optimproblem;
x = optimvar('x', 2, 'LowerBound', 0);
h = [-1, -0.15; -0.15, -2];
f=[98; 277];
prob.Objective = x'* h * x + f'*x;
A = [1, 1; 1, -2];
b=[100; 0];
prob.Constraints = A * x <= b;
[sol, fval, flag, out] = solve(prob)
sol.x          % 显示最优解
```

求得的局部最优解为：

$$x_1 = 1,\ x_2 = 1$$

目标函数的局部最优值为：371.7。

用求解非线性规划的函数 fmincon，基于求解器的程序如下。

```
clc
clear
h = [−1, −0.15; −0.15, −2];
f=[98; 277];
fx = @(x) x' * h * x+f' * x;  % 目标函数的匿名函数，x 为列向量
A = [1, 1; 1, -2];
b=[100; 0];
[x, y] = fmincon(fx, rand(2,1), A, b, [ ], [ ], [0; 0])
```

求得局部最优解为：

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 1.0 \times 10^{-7} * \begin{pmatrix} 0.5513 \\ 0.8861 \end{pmatrix}$$

目标函数的局部最优值为：2.9948e-05。

说明：如果二次规划不是凸规划，一般不用基于问题的求解方法，常用基于求解器的方法，即调用函数 fmincon。