MEMOIRS OF THE SCHOOL OF ENGINEERING, OKAYAMA UNIVERSITY, Vol. 2, No. 1, APRIL, 1967

Optimal Design of Chemical Process by Nonlinear Programming Technique*

Hayatoshi SAYAMA and Ko OI

Department of Industrial Science, School of Engineering, Okayama University, Okayama, Japan

(Received December 28, 1966)

The optimum temperature in a sequence of two stirred tanks is considered subject to inequality constraints. For an illustrative example this nonlinear programming problem is solved using the SUMT method of Fiacco and McCormick, which transforms the constrained problem into a sequence of unconstrained minimization problems. The results are presented for several cases, and are fairly good.

§ 1. Introduction

Various methods can be used to obtain the solution of optimization problems in chemical engineering. The powerful ones are dynamic programming, maximum principle, and gradient methods. Mathematical programming techniques have been widely applied to the problems in operations research. However, these techniques have not been used as actively as other optimization methods mentioned above in engineering systems.

Recently Fiacco and McCormick have presented the SUMT method[1)2)] (sequential unconstrained minimization technique) for solving convex programming problem. This method can handle problems with equality and inequality constraints and has advantage of being able to use a recent method of a "second order" gradient technique.[3)4)]

The paper shows the use of the SUMT method in determining the optimum temperature in a sequence of two stirred tanks subject to control variable inequality constraints. Numerical solutions were obtained via first and second order gradient methods.

§ 2. The Sequential Unconstrained Minimization Technique

The mathematical programming problem under consideration is to determine a vector $\bar{x}$ that solves:

(A) Primal Problem:

$$\text{minimize} \quad f(x), \tag{1}$$

subject to

$$g_i(x) \geqq 0, \quad i=1,\ 2,\ \cdots,\ m. \tag{2}$$

Here $x=(x_1,\ \cdots,\ x_n)^T$ is an n dimentional column vector, where T denotes transposition. Carroll proposed an idea for transforming the mathematical programming problem into a sequence of unconstrained minimization problems.[5)] Fiacco and McCormick extended the method of Carroll and gave the theoretical validation of the sequential unconstrained minimization technique for solving convex programming problem.[1)2)]

Denfie the function

$$P(x,r)=f(x)+r\sum_{i=1}^{m}\frac{1}{g_i(x)}. \tag{3}$$

The procedure is based on the minimizations of $P(x, r)$ over x satisfying the constraints $g_i(x)>0$, $i=1,\ 2, \cdots,\ m$, for a sequence of r values, $r_1>r_2>\ \cdots\ >r_k>\ \cdots\ >0$. Under certain conditions, there exist the minima of $P(x, r)$ represented by $x(r_1)$, $x(r_2)$, $\cdots$, $x(r_k)$, $\cdots$, and it follows that $x(r_k)\rightarrow\bar{x}$, a solution of (A), and $f[x(r_k)]\rightarrow f(\bar{x})=v_0$, the optimal solution value of (A), as $r_k\rightarrow 0$ $(k\rightarrow\infty)$.

Define $R^\circ=\{x\,|\,g_i(x)>0,\ i=1,\ \cdots,\ m\}$ and $R=\{x\,|\,g_i(x)\geqq 0,\ i=1,\ \cdots,\ m\}$. The sufficient conditions for method to accomplish (A) are as follows:

C 1: R° is nonempty.

C 2: $f(x)$ and $-g_i(x)$, $i=1, \cdots, m$, are convex and twice continuously differentiable for $x\in R$.

C 3: For every finite k, $\{x\,|\,f(x)\leqq k;\ x\in R\}$ is a bounded set.

C 4: For every $r>0$, $P(x, r_k)$ is strictly convex.

The two conditions required for any useful results are C 1 and C 3.

When C 1 to C 4 are satisfied, there is a dual

* Paper presented at the 32nd Annual Meeting of the Society of Chemical Engineers, Japan, at the Chuō University, Tokyo, April 3, 1967.

problem formulated by Wolfe.[6]

(B) Dual Problem:

$$\text{maximize} \quad G(x, u)=f(x)-\sum_{i=1}^{m} u_i \cdot g_i(x), \quad (4)$$

$$\text{subject to} \quad \nabla_x G(x, u)=0, \ u_i \geqq 0, \quad i=1, \cdots, m. \quad (5)$$

At each P-minimum the conditions $\nabla_x P[x(r_k), r_k]=0$ must hold. Letting $u_i(r_k)=r_k/g_i^2[x(r_k)]$ for $i=1, \cdots, m$, then $[x(r_k), u(r_k)]$ is a dual feasible point, and $G[x(r_k), u(r_k)] \to v_0$, the optimal solution value of (A), as $r_k \to 0$. Since v_0 is the maximum value of $G(x, u)$ for dual-feasible points, we have the following inequalities at each P-minimum,

$$G[x(r_k), u(r_k)] \leqq v_0 \leqq f[x(r_k)]. \quad (6)$$

Initial value and reduction of r: The total number of iterations required to compute the solu'ion depends on the initial value of r, r_1.

Criterion 1; Choose $r_1>0$ that minimizes the magnitude of the square of the gradient of $P(x, r)$ at x^0, the specified interior point, *i. e.*,

$$|P(x^0, r_1)|^2=\min_r |\nabla f(x^0) - r\sum_{i=1}^{m} \nabla g_i(x^0)/g_i^2(x^0)|^2.$$

Let $p(x)=\sum_{i=1}^{m} 1/g_i(x)$. Then, r_1 is given by

$$r_1=-\nabla f(x^0)^T \nabla p(x^0)/|\nabla p(x^0)|^2, \quad (7)$$

Criterion 2; Let the Hessian matrices of $f(x)$ and $p(x)$, evaluated at x^0, be denoted by H_1, H_2. One estimate of the amount by which $P(x, r)$ exceeds its minimum value is given by

$$\nabla p(x^0, r)^T [H_1+rH_2]^{-1} \nabla p(x^0, r)/2.$$

If H_1 matrix is assumed to be unimportant, then

$$r_1=\left(\frac{\nabla f(x^0)^T H_2^{-1} \nabla f(x^0)}{\nabla p(x^0)^T H_2^{-1} \nabla p(x^0)}\right)^{1/2}. \quad (8)$$

The choice of r for the $(i+1)^{st}$ minimization is given by $r_{i+1}=r_i/c$ where $c>1$.

Gradient methods: The techniques used to minimize $P(x, r)$ for various values of r are first and second order gradient method as follows:

$$x^2=x^1-\theta \nabla P(x^1), \quad (9)$$

and

$$x^2=x^1-\theta [\partial^2 P(x^1)/\partial x_i \partial x_j]^{-1} \nabla P(x^1). \quad (10)$$

The following criteria are used to terminate convergence to the minimum.

first order gradient method;

$$|\nabla P(x^i, r)|<\varepsilon, \ \varepsilon>0. \quad (11)$$

second order gradient method;

$$\nabla P(x^i, r)^T H^{-1} \nabla P(x^i, r)<\varepsilon, \quad \varepsilon>0. \quad (12)$$

where H is the Hessian matrix of P evaluated at x^i.

Final convergence criteria: The theoretical optimum value v_0 is bounded by the dual and primal function values, respectively $G[x(r), u(r)]$ and $f[x(r)]$; that is,

$$G[x(r), u(r)] \leqq v_0 \leqq f[x(r)]. \quad (13)$$

Assuming $f(\bar{x})$, v_0, and $G[x(r), u(r)]$ have the same sign, and $G[x(r), u(r)]$, $v_0 \neq 0$, and rearranging eq. (13)

$$\left|\frac{f(\bar{x})-G[x(r), u(r)]}{G[x(r), u(r)]}\right| \geqq \left|\frac{f(\bar{x})-v_0}{v_0}\right| \geqq 0 \quad (14)$$

Computational procedure:

i) Select a point x^0 interior to the feasible region.
ii) Select r_1, the initial value of r, by (7) or (8).
iii) Determine minimum of $P(x, r_k)$ for current value of r_k by gradient methods.
iv) Terminate computations if final convergence criteria are satisfied. If not, go to step v).
v) Select $r_{k+1}=r_k/c$ where $c>1$, and continue procedure from step iii).

§ 3. Process Model and Statement of Problem

For an illustrative example, the following problem was solved using the SUMT method outlined. The reaction $A \to B \to C$ is carried out in a sequence of two stirred tanks of equal residence time, τ. The reaction $A \to B$ is second order and $B \to C$ is first order, where C is the waste product. The problem is to determine the temperature T_n such that the value of $b_2+\rho a_2$ is a maximum where a_n and b_n represent the concentration of A and B at stage n and ρ is the relative value of A and B. The problem has been solved by Denn and Aris using the first order gradient method.[7)8)9)] The temperature has an inequality constraint such as

$$T^* \geqq T_n \geqq T_* \quad (15)$$

The process equations at each stage are given by

$$a_{n-1}-a_n-\tau k_1 a_n^2=0 \quad (16)$$

$$b_{n-1}-b_n+\tau k_1 a_n^2-\tau k_2 b_n=0 \quad (17)$$

where k_1 and k_2 are the rate constants for the reaction $A \to B$ and $B \to C$,

$$k_1=A_1\exp(-E_1/RT_n),$$

$$k_2=A_2\exp(-E_2/RT_n) \quad (18)$$

The numerical values used are,

$A_1=5\times10^{10}$ liters mole^{-1} min.$^{-1}$, $A_2=3.33\times10^{10}$ min.$^{-1}$ $E_1=18$ kcal, $E_2=30$ kcal, $\rho=0.3$, $N\tau=$

6 min. $a_0=1$ mole liter^{-1}, $b_0=0$ mole liter^{-1}, $T^*=370°K$, $T_*=346°K$.

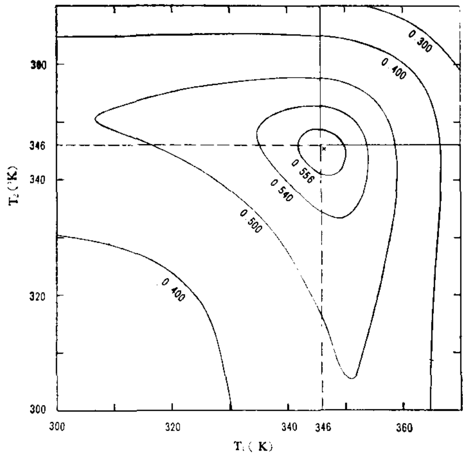

Fig. 1 Isovalue Contour Lines of Profit Function, $b_2+\rho a_2$.

Several isovalue contours of $b_2+\rho a_2$ are plotted in Fig. 1 for this set of numerical values. Fig. 1 shows that the surface of $b_2+\rho a_2$ is concave in the feasible region and in the neibourhood of maximum but is not always concave except feasible region. Optimum temperatures will be estimated from Fig. 1 that T_1 is near 346.5°K and $T_2=346.0°K$ in the feasible region. For this simple problem, the trajectories obtained by first and second order gradient methods can be plotted on T_1-T_2 plane and it is convenient to compare the convergence characteristics of these methods. The SUMT method can handle nonlinear programming problems that have a fairly large number of variables and inequality constraints.

§ 4. Result and Discussion

In the following numerical example, the methods used to minimize $P(x, r)$ for each r value are first, and second order gradient methods, eq. (9) and eq. (10), respectively. There are various computational algorithms for adjusting scale factor, θ, in these methods. The selection of θ currently used is an approximate method analogous to a binary scale factor search. Starting with an arbitrary $\theta(\theta=100)$, this method requires determination of an integer, n, which locally minimizes the expressions

$$P[x^1-2^{-n}\theta\ \nabla P(x^1),\ r]$$

or

$$P[x^1-2^{-n}\theta(\partial^2 P(x^1)/\partial x_i \partial x_j)^{-1}\ \nabla P(x^1),\ r]$$

For this n value, the second point, x^2, is obtained by choosing the point which gives the smaller values of P function for $2^{-n}\theta$ and (3/2) $2^{-n}\theta$. θ value at the second point is obtained from either $2^{-n}\theta$ or (3/2) $2^{-n}\theta$ at the starting point. The procedure is repeated until the minimum is approached. The criterion used to terminate convergence to the minimum is eq. (11), $|\nabla P(x^i, r)|<\varepsilon=10^{-5}$, to compare the number of iterations for both methods on the same bases.

The initial value of r, r_1, computed by Criterion 1, eq. (7), may be negative for some point in the feasible region, and Criterion 2, eq. (8), is used. c value for reduction of r is taken as $20=r_i/r_{i+1}$.

Profit function $-(b_2+\rho a_2)$ is minimized, as the SUMT method is formulated to obtain the minimum.

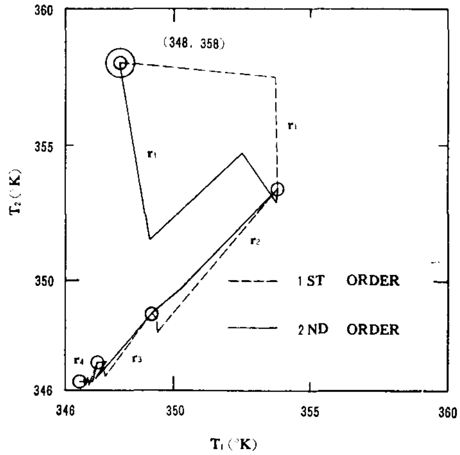

Fig. 2 Iteration Trajectories for First and Second Order Gradient Method.

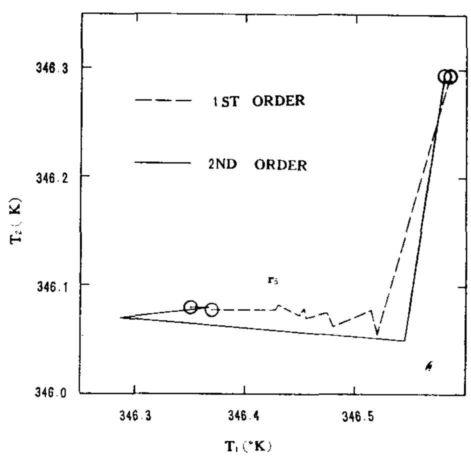

Fig. 3 Iteration Trajectories for First and Second Order Gradient Method.

Table 1. Computer Solution by First Order Gradient Method.

	r	*	**	$\|\nabla P\| \times 10^{+5}$	P	f	G	$\|(f-G)/G\|$	T_1	T_2
1	0.468	5	5	0.92	−0.332446	−0.512508	−0.692571	0.26	353.858	353.392
2	0.234×10^{-1}	5	10	0.31	−0.534780	−0.552288	−0.569796	0.31×10^{-1}	349.280	348.877
3	0.117×10^{-2}	6	16	0.31	−0.556560	−0.558737	−0.560914	0.39×10^{-2}	347.319	346.985
4	0.585×10^{-4}	9	25	0.20	−0.559387	−0.559690	−0.559993	0.54×10^{-3}	346.584	346.296
5	0.293×10^{-5}	10	35	0.91	−0.559794	−0.559841	−0.559888	0.83×10^{-4}	346.370	346.076

*Moves Required to Minimize P(x, r). **Cumulative Number of Moves

Table 2. Computer Solution by Second Order Gradient Method.

	r	*	**	$\|\nabla P\| \times 10^{+5}$	P	f	G	$\|(f-G)/G\|$	T_1	T_2
1	0.468	6	6	0.47	−0.332446	−0.512477	−0.692508	0.26	353.859	353.396
2	0.234×10^{-1}	4	10	0.32	−0.534780	−0.552289	−0.569799	0.31×10^{-1}	349.279	348.877
3	0.117×10^{-2}	4	14	0.21	−0.556560	−0.558738	−0.560915	0.39×10^{-2}	347.316	346.986
4	0.585×10^{-4}	4	18	0.30	−0.559387	−0.559691	−0.559995	0.54×10^{-3}	346.582	346.295
5	0.293×10^{-5}	4	22	0.67	−0.559795	−0.559842	−0.559889	0.84×10^{-4}	346.352	346.076
6	0.146×10^{-6}	3	25	0.56	−0.559861	−0.559870	−0.559879	0.16×10^{-4}	346.330	346.018
7	0.731×10^{-8}	2	27	0.51	−0.559875	−0.559877	−0.559878	0.33×10^{-5}	346.332	346.004
8	0.366×10^{-9}	2	29	0.60	−0.559878	−0.559878	−0.559878	0.74×10^{-6}	346.332	346.001

Results for successve approximations of temperature obtained by first and second order gradient methods are shown in Figs. 2 and 3. The initial starting point is chosen at (T_1, T_2) =(348°K, 358°K). Tables 1 and 2 give the computer solution to the problem. It is shown from these figures and tables that second order gradient method has more efficient convergence characteristics than first order method. The step size is so small at the step of r_6 that round-off errors make it difficult to obtain correct solution values via first order gradient method. Second order gradient method can continue further minimizations.

Table 3 shows the number of iterations against starting points in the feasible region. It can be seen from Tables 1~3 that second order gradient method requires about 20 number of iterations for convergence at the step of r_5, where first order method requires about 40 number of iterations starting from the same point. Values of P, f and G taken from Table 2 are plotted against the number of iterations in Fig. 4. Convergence to the solution values is fairly good by second order gradient method. Table 2 shows that five and seven place profit function agreements are achieved at the step of r_5 and r_8, respectively.

Table 3. Number of Iterations vs. Initial Starting Points.

T_1 (°K)	358		368	
T_2 (°K)	348		368	
r_1	0.397		0.0378	
Gradient Method	1ST	2ND	1ST	2ND
1	6	6	5	4
2	5	4	5	5
3	6	3	6	4
4	9	3	17	4
5	22	3	13	3
TOTAL	48	19	46	20

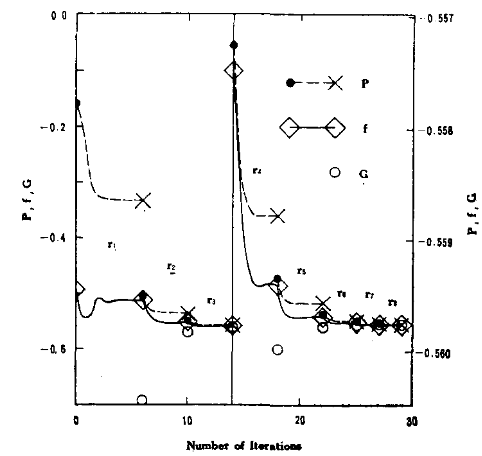

Fig. 4 Successive Approximations to the Optimal Solution Value.

References

1) A. V. FIACCO and G. P. MCCORMICK: Management Science **10** (1964) 360.
2) A. V. FIACCO and G. P. MCCORMICK: Management Science **10** (1964) 601.
3) W. C. DAVIDON: A. E. C. Res. Dev. Rept. **ANL-5990** (November 1959).
4) R. FLETCHER and M. J. D. POWELL: British Computer J. **6** (1963) 163.
5) C. W. CARROLL: Operations Research **9** (1961) 169.
6) P. Wolfe: Quart. Appl. Math. **XIX** (1961) 239.
7) M. M. DENN and R. ARIS: Ind. Eng. Chem. Fundamentals **4** (1965) 7.
8) M. M. DENN and R. ARIS: Ind. Eng. Chem. Fundamentals **4** (1965) 213.
9) M. M. DENN and R. Aris: Ind. Eng. Chem. Fundamentals **4** (1965) 248.