

一种可同时提高寻优速度和寻优精度的寻优方案*

汪 成 为

摘 要

本文提出了一种可同时提高寻优速度和寻优精度的寻优方案，并给出了用这种寻优方案求解实际问题的一个例题。

一、概 述

自动寻优器是一种专用计算装置，它可以作为数学计算机及控制计算机使用，也可作为某些自适应系统的反馈元件，当与通用模拟计算机联用时又可完成自动逼近、自动综合等任务^[2]。

简言之，自动寻优器的职能是使系统（或被自动寻优器所控制的对象）在满足一定的限制条件(H_0)下，使质量指标(Q)达极值。

$$Q = Q(x_1, x_2, \dots, x_n) \max \text{ 或 } \min, \tag{1}$$
$$H_j = H_j(x_1, x_2, \dots, x_n) \leq 0, \quad j = 1, 2, \dots, m,$$

则限制条件为

$$H_0 = \sum_1^m H_j \leq 0. \tag{2}$$

为了在设计及使用自动寻优器时更为方便，我们不直接把 Q 函数输入到自动寻优器，而是输入 ε 函数，其关系为

$$\varepsilon = K|Q - Q_0|, \tag{3}$$

其中 $Q > 0$ ， Q_0 是给定的或企望的质量指标。经过这样一个转换后，自动寻优器的职能就成为在满足已给定的限制条件 H_0 的前提下，寻找能使 ε 函数为极小的各变量 x_i 的值

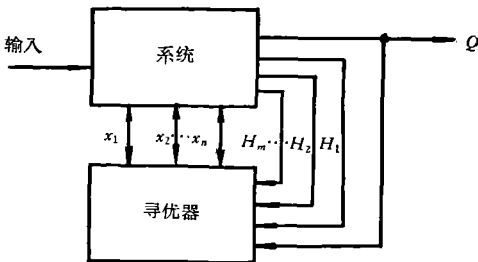


图 1

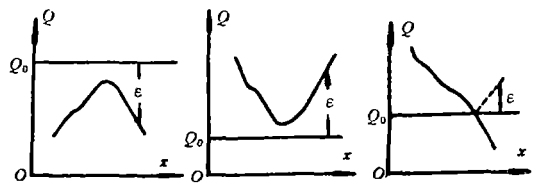


图 2

* 本文曾于1965年7月在中国自动化学会年会上宣读。本文于1965年2月13日收到。

($i = 1, 2, \dots, n$).

梯度法与最速降落法的基本原理在文献[2], [3]内均有详尽的叙述, 本文不作赘述。以下只扼要说明求函数梯度的机器方法及决定工作步程大小的各种因素。

如在完成第 p 步工作步程时的质量函数为 $\varepsilon_p(x_{1p}, x_{2p}, \dots, x_{np})$, 则在进行第 $p + 1$ 步工作步程前应首先进行试探步程, 即在机器上求出梯度 $\nabla \varepsilon_p$, 即要求出

$$\nabla \varepsilon_p = i_1 \frac{\partial \varepsilon_p}{\partial x_1} + i_2 \frac{\partial \varepsilon_p}{\partial x_2} + \dots + i_i \frac{\partial \varepsilon_p}{\partial x_i} + \dots + i_n \frac{\partial \varepsilon_p}{\partial x_n}$$

内的各 $\frac{\partial \varepsilon_p}{\partial x_i}$ 量 ($i = 1, 2, \dots, n$).

机器内求 $\frac{\partial \varepsilon_p}{\partial x_i}$ 的方法为: 令 Δx_ε 为一增量, 且认为

$$\varepsilon_{ip} = \varepsilon(x_{1p}, x_{2p}, \dots, x_{ip}, \dots, x_{np}), \quad (4)$$

$$\varepsilon'_{ip} = \varepsilon(x_{1p}, x_{2p}, \dots, x_{ip} + \Delta x_\varepsilon, \dots, x_{np}), \quad (5)$$

当 Δx_ε 足够小时, 可近似地认为

$$\Delta \varepsilon_{ip} = \varepsilon'_{ip} - \varepsilon_{ip} = k \frac{\partial \varepsilon_p}{\partial x_i}, \quad k = \text{常量}. \quad (6)$$

在完成第 $p + 1$ 步工作步程时得 $x_{i(p+1)}$,

$$x_{i(p+1)} = x_{ip} + \int^{\Delta \tau} k \frac{\partial \varepsilon_p}{\partial x_i} d\tau = x_{ip} + C \Delta \varepsilon_{ip}. \quad (7)$$

因为欲求 ε_{\min} 值, 故常数 C 小于零。

由(7)式可知, 工作步程是沿梯度的反方向运动(因为 $C < 0$), 其步程的大小决定于 Δx_ε , C 及 ε 函数的性质。

在完成第 $p + 1$ 步工作步程时的质量指标为

$$\varepsilon_{(p+1)} = \varepsilon(x_{1(p+1)}, x_{2(p+1)}, \dots, x_{n(p+1)}).$$

当寻优轨迹破坏了限制条件而进入非允许区时, 则可用上述机器求梯度的方法对 H 函数求梯度, 然后按梯度的反方向运动, 使寻优轨迹又回到允许区内。增量 Δx_H 可等于 Δx_ε , 也可不等。以下只讨论在允许区内的寻优过程。

二、采用一般梯度法时其寻优速度与寻优精度间的矛盾

当自动寻优器的具体线路及质量函数已确定, 并且寻优初始点也已决定时, 则为了加速寻优过程就必须加大每一步工作步程, 由(5), (6), (7)式可知, 也即是增大 Δx_ε 与 C 值。但是, 当 Δx_ε 增大时, 由机器所求出的梯度与理论梯度在方向上的偏离也随之增大, 这样就使实际的寻优轨迹与理论轨迹发生很大的偏离, 用梯度法寻优时其理论轨迹应恆垂直于 ε 函数的等值面。

增大 C 值也可增大工作步程。但是, 由于机器所求出的梯度与理论梯度在方向上存在偏离, 以及在一步工作步程的范围内理论梯度是变化的, 所以 C 值的增大也使实际的寻优轨迹与理论轨迹的偏离加大。

以下, 再讨论 Δx_ε 和 C 值与最终(寻找到极值 ε_{\min} 时)寻优精度的关系。

为讨论方便起见, 我们假定 $\varepsilon = Ax^2$, 则显然^[1]

$$\left. \begin{aligned} \Delta \varepsilon_p &= A \Delta x_\varepsilon (-2x_p + \Delta x_\varepsilon) \quad (\text{如 } x_0 > 0) \\ \Delta \varepsilon_p &= A \Delta x_\varepsilon (2x_p + \Delta x_\varepsilon) \quad (\text{如 } x_0 < 0), \end{aligned} \right\} \quad (8)$$

且只有在最终(第 ∞ 步)当

$$\left. \begin{aligned} x_\infty &= \frac{1}{2} \Delta x_\varepsilon \quad (\text{如 } x_0 > 0) \\ \text{或 } x_\infty &= -\frac{1}{2} \Delta x_\varepsilon \quad (\text{如 } x_0 < 0) \end{aligned} \right\} \quad (9)$$

时 $\Delta \varepsilon$ 才可能等于零。因此 $\frac{1}{2} \Delta x_\varepsilon$ 点或 $-\frac{1}{2} \Delta x_\varepsilon$ 点是理论上的稳定状态。因为理想的极值点为 $x=0$ ，因此，稳定运动的幅值(或称为寻优误差)为 $\frac{1}{2} \Delta x_\varepsilon$ ，试探运动的增量(Δx_ε)愈小，寻优精度愈高。

但是，实际上由于自动寻优器无法无限地提高对 $\Delta \varepsilon$ 的敏感性，同时又因为自动寻优器线路本身存在杂音与漂移，外界还输入各种扰动，因此，为使自动寻优器可靠地工作也必须规定对 $\Delta \varepsilon$ 的极限敏感值，以下称为 $\overline{\Delta \varepsilon}$ 。

在图 3 中可看到，相应于 $\overline{\Delta \varepsilon}$ 可得

$$\begin{aligned} x &= x' \quad (\text{如 } x_0 < 0), \\ x &= x'' \quad (\text{如 } x_0 > 0), \end{aligned}$$

且

$$|x'| = |x''|.$$

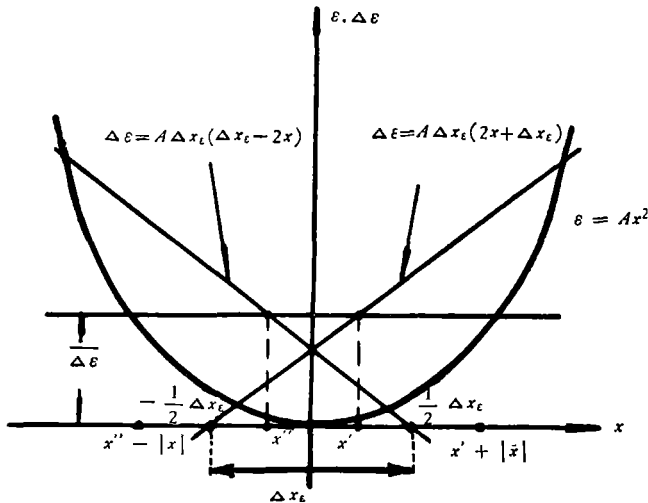


图 3

因此，自动寻优器最终的稳定状态是 x 在 $(x'' - |\bar{x}|, x' + |\bar{x}|)$ 内振荡，而

$$|\bar{x}| = C \overline{\Delta \varepsilon}.$$

因此，自动寻优器稳定运动后与零点的偏差范围 l 为

$$l = (x' + |\bar{x}|). \quad (10)$$

以上是对 $\varepsilon = Ax^2$ 形式的 ε 函数进行讨论，但并不因此而使所得的结论在概念上没

有广泛的意义。当

$$\varepsilon = \sum_{i,j=1}^n A_{ij}x_i x_j$$

时,我们也可进行同样的讨论。

可见,在采用一般文献内所介绍的梯度法(或基于梯度法基本原理而发展成的最速降落法)寻优时,无法同时提高寻优速度和寻优精度。通常是从对寻优结果的要求出发,决定是争取加速寻优速度,或是致力于提高寻优精度,然后根据我们对系统(或被自动寻优器所控制的对象)的了解,对 ε 函数的了解,或通过试解来选择较理想的寻优初始点、 Δx_ε (或 Δx_H) 值、 C 值,或调整 $\overline{\Delta\varepsilon}$ 值的大小。但这往往增加了工作困难,减低了工作效率。

三、采用非线性系数 K_ε (或 K_H) 的方案

在上述梯度法的基本原理上,我们引入了非线性系数 K_ε (或 K_H),从而可能同时提高寻优速度与寻优精度。

当把 ε 函数(或 H 函数)由数学方程编排成被模拟计算机所运算的机器方程时,必须引入比例系数 K_ε (或 K_H),即

$$\left. \begin{aligned} \varepsilon_m &= K_\varepsilon \cdot \varepsilon, \\ H_m &= K_H \cdot H, \end{aligned} \right\} \quad (11)$$

脚注 m 表示机器量, K_ε 与 K_H 为常量。

现在,我们使 K_ε (或 K_H) 成为 ε (或 H) 的单调下降函数,即当 ε (或 H) 增大时 K_ε (或 K_H) 减小。在此必须强调指出,这种非线性关系只是为了自动寻优器的需要而成立,真正的质量指标或限制条件仍为式(1),(2)所定义。

当自动寻优器结束了第 $p-1$ 步工作步程并开始进行求梯度前,自动寻优器的逻辑控制线路应使非线性系数按已定的关系随第 $p-1$ 步工作步程所得到的 ε 值(或 H 值)而变化,即

$$\left. \begin{aligned} K_{\varepsilon p} &= K_{\varepsilon p}(\varepsilon_{p-1}) \\ K_{H p} &= K_{H p}(H_{p-1}) \end{aligned} \right\}, \quad (12)$$

则

$$\left. \begin{aligned} \varepsilon_{p m}^* &= K_{\varepsilon p}(\varepsilon_{p-1}) \cdot \varepsilon_p \\ \text{或} &= \frac{K_{\varepsilon p}(\varepsilon_{p-1})}{K_\varepsilon} \cdot \varepsilon_{p m} \\ H_{p m}^* &= K_{H p}(H_{p-1}) \cdot H_p \\ \text{或} &= \frac{K_{H p}(H_{p-1})}{K_H} \cdot H_{p m} \end{aligned} \right\}, \quad (13)$$

其中 $\varepsilon_{p m}^*$ 与 $H_{p m}^*$ 是输入到自动寻优器的机器量,自动寻优器将对 $\varepsilon_{p m}^*$ (或 $H_{p m}^*$) 求梯度与完成第 p 步工作步程。上述原理可用图 4 所示的框图说明之。

由于采用了非线性系数,所以,虽然当 ε_p (或 H_p) 值逐渐趋小,输入到自动寻优器的 $\varepsilon_{p m}^*$ (或 $H_{p m}^*$) 始终比原 $\varepsilon_{p m}$ (或 $H_{p m}$) 值大。显然,当 ε (或 H) 值为零时, ε_m^* (或 H_m^*) 亦为零。

由式(8)可知,当 Δx_ε 已定时,则偏差方程

$$\Delta \varepsilon_p = A \Delta x_\varepsilon (\pm 2x_p + \Delta x_\varepsilon)$$

也被确定。以下我们将写出引入非线性系数后的偏差方程。为简便起见，我们假定 x 量的机器比例系数等于 1 (因此对 x 量不必加脚注 m)。

$$\Delta \varepsilon_{pm}^* = AK_{\varepsilon_p}(\varepsilon_{p-1}) \cdot \Delta x_\varepsilon (\pm 2x_p + \Delta x_\varepsilon), \quad (14)$$

而原式(8)的机器方程为

$$\Delta \varepsilon_{pm} = AK_\varepsilon \Delta x_\varepsilon (\pm 2x_p + \Delta x_\varepsilon), \quad (15)$$

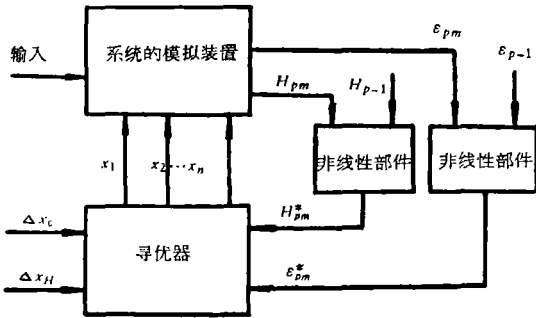


图 4

如 $K_{\varepsilon_p}(\varepsilon_{p-1}) > K_\varepsilon$ ，则对相同的 Δx_ε 与 x_p 而言必然 $\Delta \varepsilon_{pm}^* > \Delta \varepsilon_{pm}$ 。因此，在选择同样的 Δx_ε 与 C 值的情况下，采用非线性系数后可加速寻优过程。事实上当寻优轨迹愈接近极值时， $K_{\varepsilon_p}(\varepsilon_{p-1})$ 愈大于 K_ε ，同时可选取适当的非线性系数的形状，使在寻优过程开始时 $K_{\varepsilon_p}(\varepsilon_{p-1})$ 不小于(或等于) K_ε 值。

对于第 p 步工作步程而言，如果对两种方法所选取的 C 值相同，则在采用非线性系数后，可减小试探增量 Δx_ε 而仍能得到同样大小的工作步程。因此在引入非线性系数时，可适当减小 Δx_ε (或 Δx_H) 值，从而使寻优轨迹与理论轨迹更为吻合。

以下将说明引入非线性系数后可提高最终(寻找到 ε_{\min} 时)寻优精度的原因。

如果自动寻优器对偏差量的极限敏感值仍为 $\overline{\Delta \varepsilon}$ ，则在未引入非线性系数前其偏差量方程为式(15)所示，相应于 $\overline{\Delta \varepsilon}$ 可得

$$x' = \frac{\overline{\Delta \varepsilon}}{2AK_\varepsilon \Delta x_\varepsilon} - \frac{\Delta x_\varepsilon}{2}, \quad (16)$$

$$x'' = \frac{\Delta x_\varepsilon}{2} - \frac{\overline{\Delta \varepsilon}}{2AK_\varepsilon \Delta x_\varepsilon}. \quad (17)$$

令 x' 点距 $-\frac{1}{2} \Delta x_\varepsilon$ 点的距离为 $l_{x'}$ ， x'' 点距 $\frac{1}{2} \Delta x_\varepsilon$ 点的距离为 $l_{x''}$ ，可得

$$l_{x'} = l_{x''} = \frac{\overline{\Delta \varepsilon}}{2AK_\varepsilon \Delta x_\varepsilon}. \quad (18)$$

当引入非线性系数时，其偏差方程为式(14)所示，相应于 $\overline{\Delta \varepsilon}$ 可得

$$x'^* = \frac{\overline{\Delta \varepsilon}}{2AK_{\varepsilon_s}(\varepsilon_{s-1}) \Delta x_\varepsilon} - \frac{\Delta x_\varepsilon}{2}, \quad (19)$$

$$x''^* = \frac{\Delta x_\varepsilon}{2} - \frac{\overline{\Delta \varepsilon}}{2AK_{\varepsilon_s}(\varepsilon_{s-1}) \Delta x_\varepsilon}, \quad (20)$$

脚注 s 是表示在完成第 s 步工作步程时能使 $\Delta \varepsilon_{sm}^* = \overline{\Delta \varepsilon}$ 。令 x'^* 点距 $-\frac{1}{2} \Delta x_\varepsilon$ 的距离为 $l_{x'^*}$ ， x''^* 点距 $\frac{1}{2} \Delta x_\varepsilon$ 点的距离为 $l_{x''^*}$ ，则可得

$$l_{x'^*} = l_{x''^*} = \frac{\overline{\Delta \varepsilon}}{2AK_{\varepsilon_s}(\varepsilon_{s-1}) \Delta x_\varepsilon}. \quad (21)$$

通常情况下 $K_{\epsilon_s}(\epsilon_{s-1}) \gg K_{\epsilon}$, 所以 $l_{x'}^* \ll l_{x'}$. 因此, 如式(10)所示, 在未引入非线性系数时, 自动寻优器稳定运动后与零点的偏差范围为

$$l = \left(\frac{\overline{\Delta \epsilon}}{2AK_{\epsilon}\Delta x_{\epsilon}} - \frac{\Delta x_{\epsilon}}{2} + c\overline{\Delta \epsilon} \right). \tag{22}$$

当引入非线性系数后, 稳定运动后与零点的偏差范围为

$$l^* = \left(\frac{\overline{\Delta \epsilon}}{2AK_{\epsilon_s}(\epsilon_{s-1})\Delta x_{\epsilon}} - \frac{\Delta x_{\epsilon}}{2} + c\overline{\Delta \epsilon} \right). \tag{23}$$

则可得

$$\Delta l = l - l^* = \frac{\overline{\Delta \epsilon} [K_{\epsilon_s}(\epsilon_{s-1}) - K_{\epsilon}]}{2AK_{\epsilon} \cdot K_{\epsilon_s}(\epsilon_{s-1})\Delta x_{\epsilon}} \geq 0, \tag{24}$$

可见提高了最终寻优精度。

如 $\epsilon = \frac{1}{2}x^2$, x 的寻优范围为 -100 伏至 $+100$ 伏, 为保证模拟计算机不过载, 取 $K_{\epsilon} = \frac{1}{50}$, 又假定 $\Delta x_{\epsilon} = 2$ 伏, $C = 10$, $\overline{\Delta \epsilon} = 0.1$ 伏. 如能使 $K_{\epsilon_s}(\epsilon_{s-1}) = \frac{1}{5}$, 则可得 $l = 2.5$ 伏, $l^* = 0.25$ 伏, $\Delta l = 2.25$ 伏, 大大提高了最终寻优精度。

图 5 表示出在引入非线性系数后, 在进行第 s 步工作步程时, 偏差方程与 $l_{x'}^*$, x'^* 等量的几何关系。

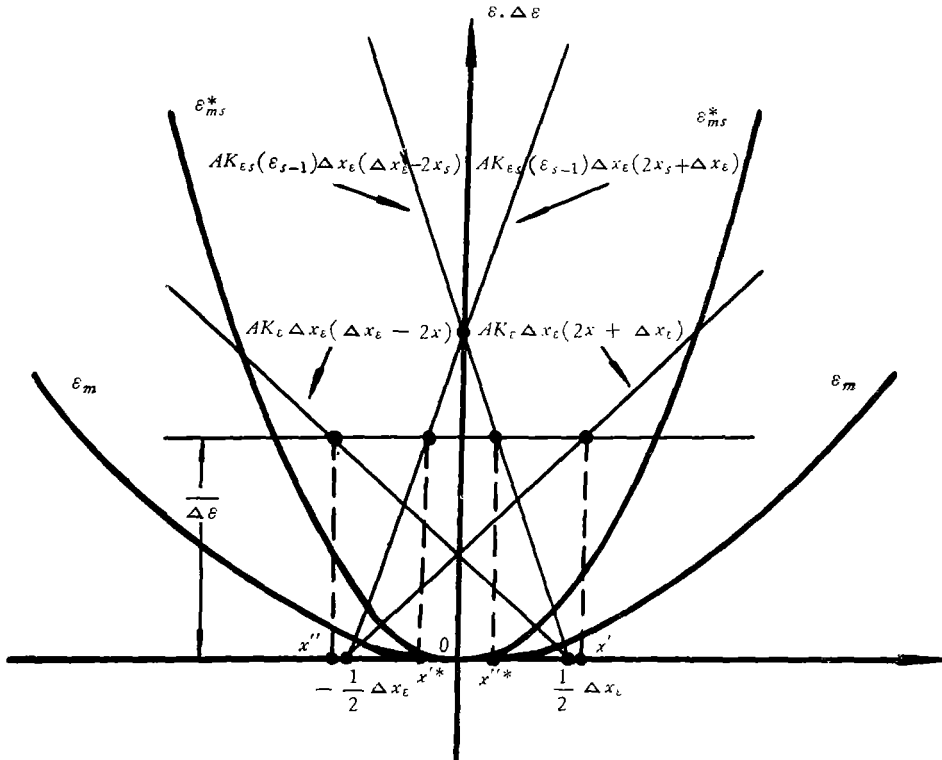


图 5

用电子式非线性部件完成。在自动寻优器的逻辑控制系统中引出 K_1 与 K_2 指令，并由它们控制存储器和电子门。

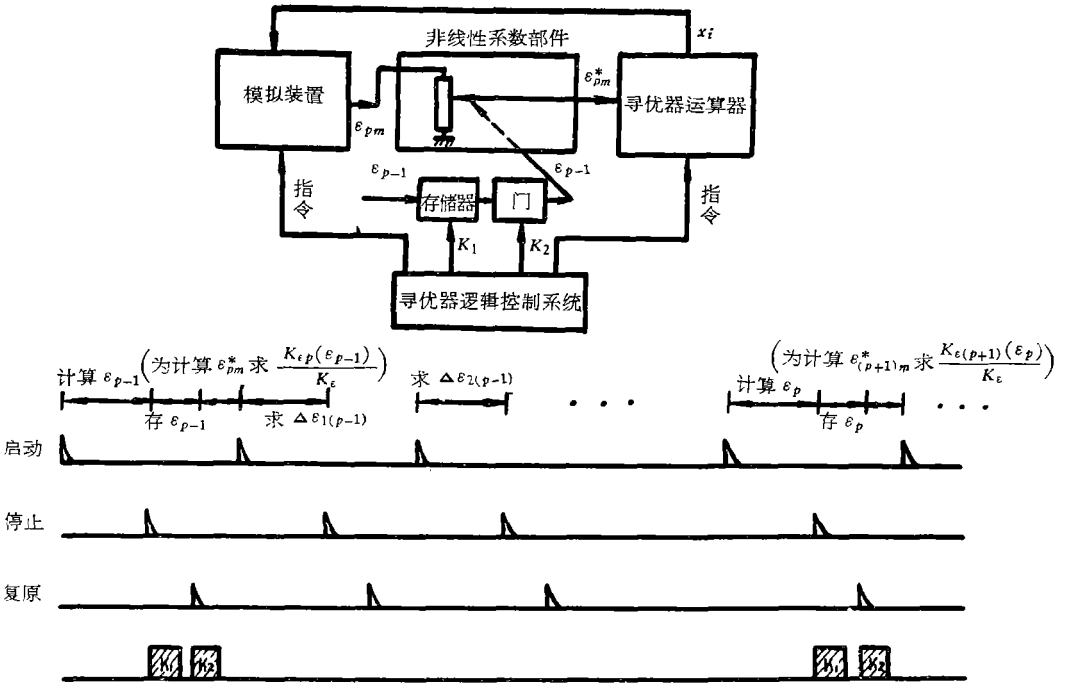


图 6

我们可以选择各种不同的单调下降曲线作为非线性系数，但必须保证在整个寻优过程中 ϵ_{pm}^* 值不会引起模拟计算机的过载，同时为了提高寻优精度，可使非线性系数在愈接近 ϵ_{p-1} 的零值时斜率愈大(如图 7 所示)。

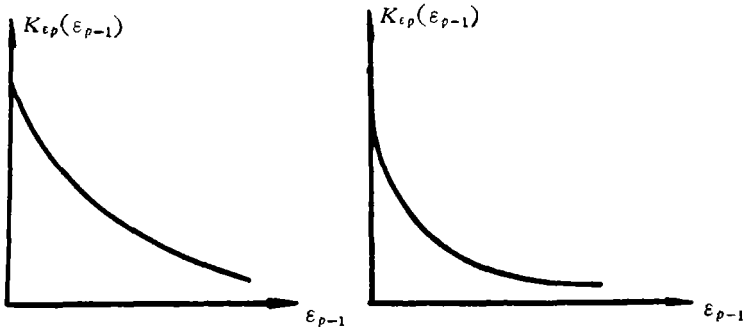


图 7

四、例 题

假定 ϵ 函数为

$$\epsilon = 0.171x_1^2 + 0.441x_1x_2 + 0.941x_2^2 \tag{25}$$

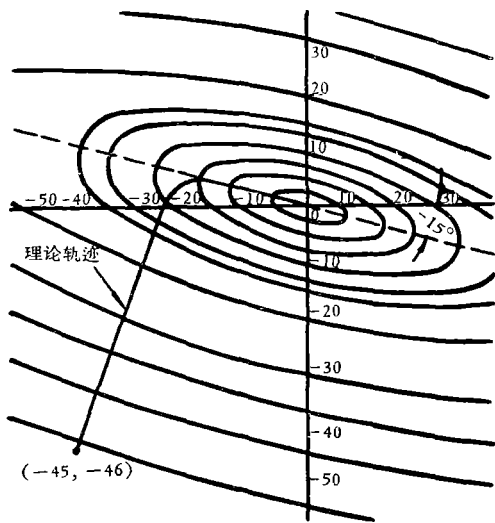


图 8

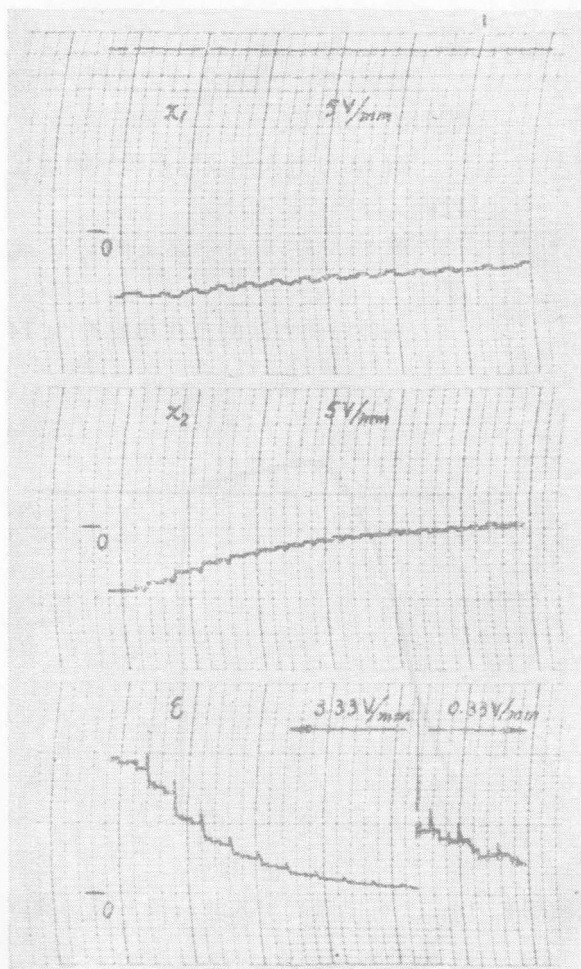


图 9

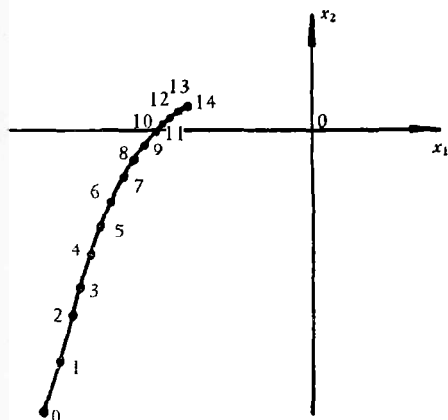


图 10

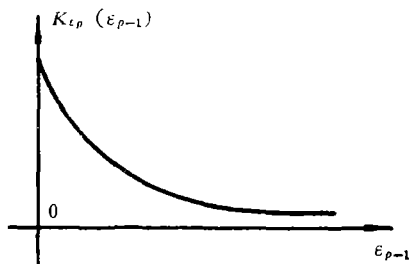


图 11

显然,这是 x_1x_2 平面上的一组椭圆曲线,其长轴 a 与短轴 b 的关系为

$$a = 3b,$$

长轴与 ox_1 轴所夹的锐角为 -15° , 椭圆中心位于坐标原点上(如图 8 所示)。显然, ε_{\min} 为点 $(0, 0)$ 。由于应用了乘法器(求 x_1^2, x_1x_2, x_2^2), (25)式改为

$$\frac{\varepsilon}{100} = 0.00171x_1^2 + 0.00444x_1x_2 + 0.00940x_2^2 \tag{26}$$

当我们选择了初始寻优点为 $(-45, -46)$ 时, 我们选定 $K_\varepsilon = \frac{1}{50}$, 这样可较满刻度地使用模拟计算机。因此, 最终可得

$$\varepsilon_m = \frac{1}{50} \varepsilon = 2(0.00171x_1^2 + 0.00444x_1x_2 + 0.00940x_2^2) \tag{27}$$

我们确定 $\Delta x_\varepsilon = 2$ 伏。

图 9 为用一般梯度法时的寻优过程。

图 10 为在 x_1x_2 平面上的寻优轨迹。

然后, 我们引入非线性系数, 其形状如图 11 所示。

图 12 为引入非线性系数后的寻优过程。

图 13 为在 x_1x_2 平面上的寻优轨迹。

两种寻优方法的工作步程均为 14

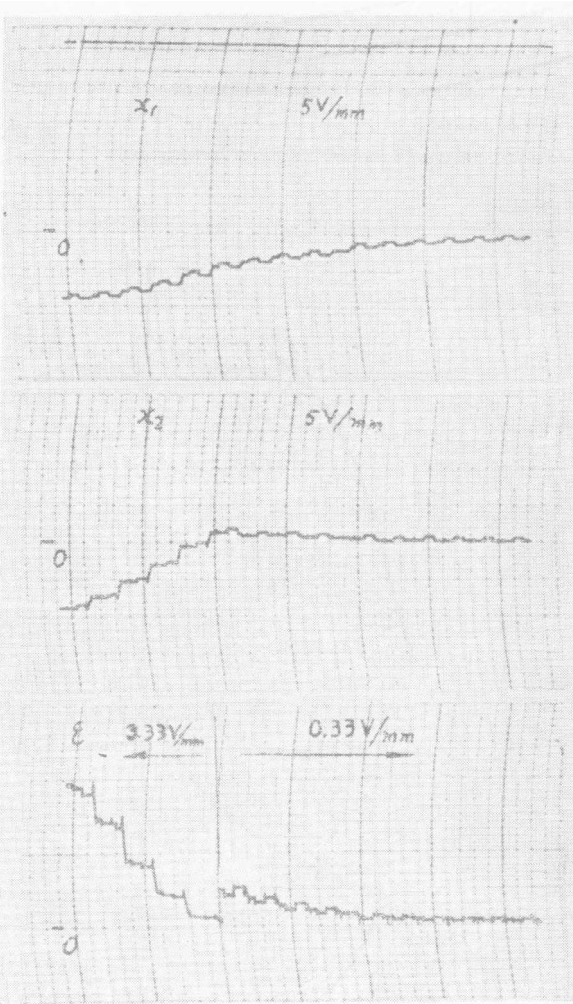


图 12

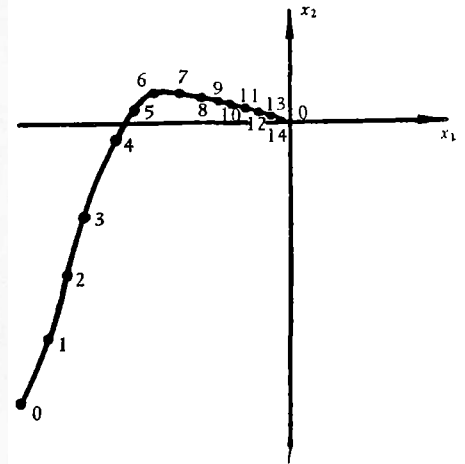


图 13

步, 但所达到的最终精度确大不相同。比较图 8, 10, 13 可以看到引入非线性系数后的寻优轨迹十分接近于理论轨迹。

范存智同志参加了上述的全部试验工作, 并提出了许多宝贵意见。

参 考 文 献

- [1] Стаховский, Р. И., О сравнении некоторых методов поиска для автоматического оптимизатора, Теория и применение дискретных автоматических систем—Труды конференции, 1958.
- [2] Фельдбаум, А. А., Автоматический оптимизатор, *AuT* (1958), № 8.
- [3] Шубин, А. Б. 10-канальный автоматический электронно-релейный оптимизатор, *AuT* (1960), № 5.

A SCHEME OF AUTOMATIC OPTIMIZATION WITH HIGHER SEARCHING SPEED AND PRECISION

WANG CHENG-WEI

By introducing a non-linear coefficient in the error function of the performance index and constraint condition, the searching speed and precision of automatic optimizer based on gradient method can be improved. An example is given to illustrate the use of this scheme.