哈尔滨商业大学

学者文库

HAERBIN SHANGYE DAXUE XUEZHE WENKU

经济管理中的优化方法

徐耀群 著

中国财政经济出版社

哈尔滨商业大学学者文库

经济管理中的优化方法

徐耀群 著

中国财政经济出版社

图书在版编目（CIP）数据

经济管理中的优化方法/徐耀群著 .—北京：中国财政经济出版社，2004.12

（哈尔滨商业大学学者文库）

ISBN 7-5005-7800-8

Ⅰ.经… Ⅱ.徐… Ⅲ.经济管理-最优化算法 Ⅳ.F224.31

中国版本图书馆 CIP 数据核字（2004）第 133088 号

中国财政经济出版社 出版

URL：http：//www.cfeph.com.cn

E-mail：cfeph @ cfeph.cn

社址：北京市海淀区阜成路甲 28 号　邮政编码：100036

发行处电话：88190406　财经书店电话：64033436

北京财经印刷厂印刷　各地新华书店经销

880×1230 毫米　32 开　8.125 印张　205 000 字

2004 年 12 月第 1 版　2004 年 12 月北京第 1 次印刷

印数：1—1 500　定价：18.00 元

ISBN 7-5005-7800-8/F·6850

（图书出现印装问题，本社负责调换）

内 容 简 介

本书系统地介绍了经济管理中的优化理论及其应用，内容包括无约束优化、线性规划、整数规划、目标规划、动态规划、非线性规划、遗传算法、神经网络优化、混沌优化。本书在编写过程中力求理论严谨，概念叙述清晰、凝炼，并配有实例介绍优化方法的应用。书中各章配有一定量习题，以供读者理解、巩固知识，锻炼解决实际问题的能力。

本书可供高等院校经济管理类专业的硕士及博士研究生作为教材使用，也可作为工程技术人员及经济管理人员学习的参考书。

前言

最优化问题是人们在工程技术、科学研究、经济管理和交通运输等领域经常遇到的问题。追求最优目标是人类的理想，也是大自然最普遍的运行规则。最优化方法就是从众多可能方案中选择最佳者，以达到最优目标的科学。优化理论是一门比较新兴的应用数学分支，在工程技术、自然科学及经济管理诸多领域中有其广泛的应用前景，已日益受到人们的重视。在有限的资源条件下如何使产出最多（或利润最大）；在市场需求一定的情况下如何节省人力、物力、财力；结构设计在满足强度的要求下如何使所用材料最轻。这些实际问题促使最优化理论产生、发展及日益繁荣，其应用已广泛渗透到社会的各个领域，即使在日常生活中也经常出现优化的身影。可以断言，随着信息社会的进一步发展，最优化理论和技术必将在社会的诸多方面起着越来越大的作用。

本书共分 9 章，系统地介绍了最优化理论的基本内容。本书在概念叙述上省去了许多详细的说明，在举例上也尽量做到凝练简洁，没有给出过多的冗繁例子；对于一些证明较冗长和复杂的定理，只给出了定理叙述，证明从略。从总体上讲，本书在内容编排上尽量做到系统、全面，重点突出，选材针对性强，具有较好的可读性。

优化理论和方法的内容极其广博，本书的最大特点是力图涵盖最优化理论的最新内容，将近年来出现的遗传算法、神经网络优化及混沌优化在最优化理论的框架中加以介绍，使读者可以了解最优化理论发展的现状。书中很多内容现在正处于研究发展阶段，也可为感兴趣的读者做入门的介绍，以便使更多的科研、工程技术及管理人员参与到优化理论的研究与开发应用上来。

由于作者水平有限，编写时间匆忙，书中难免存在不妥之处，敬请广大读者批评指正。

徐耀群

2004 年 9 月

目录

第 1 章　无约束规划

无约束规划是指没有任何约束条件的优化问题，一般地讲，在经济管理领域中的优化问题往往受资源、资金、成本等限制。当变量的范围很大时，可以把有约束问题近似看成无约束问题来处理。

1.1　基本概念

设 $f(X)$ 为定义在 n 维欧氏空间 R^n 上的 n 元实函数，我们把寻找 $f(X)$ 的极小点的问题称为一个无约束最优化问题。这个问题可用下列形式表示：

$$\min f(X),\ X=(x_1,\ x_2,\ \cdots,\ x_n)^T \in R^n \tag{1.1}$$

其中，$f(X)$ 称为目标函数。

函数的极值分为局部极值和全局极值，下面给出它们的定义。

定义 1　(1) 设 $f(X)$ 为定义在 n 维欧氏空间 R^n 的某一区域 Ω 上的 n 元实函数，其中 $X=(x_1,\ x_2,\ \cdots,\ x_n)^T$。对于 $X^* \in R$，如果存在某个 $\varepsilon>0$，使所有与 X^* 的距离小于 ε 的 $X\in\Omega$（即 $X\in\Omega$ 且 $\|X-X^*\|<\varepsilon$）均满足不等式 $f(X)\geqslant f(X^*)$，则称 X^* 为 $f(x)$ 在 R 上的局部极小点，$f(X^*)$为局部极小值。若对于所有 $X\neq X^*$ 且与 X^* 的距离小于 ε 的 $X\in\Omega$，$f(x)>f(X^*)$，则称 X^* 为 $f(x)$ 在 Ω 上的严格局部极小点，$f(X^*)$ 为严格局部极小值。

(2) 对于点 $X^*\in\Omega$，若对于所有 $X\in\Omega$ 都有 $f(X)\geqslant f(X^*)$，则称 X^* 为 $f(x)$ 在 Ω 上的全局极小点，$f(X^*)$ 为

全局极小值。若对于所有 $X\in\Omega$ 且 $X\neq X^*$，都有 $f(X)>f(X^*)$，则称 X^* 为 $f(X)$ 在 Ω 上的严格全局极小点，$f(X^*)$为严格全局极小值。

如将上述不等式反向，即可得到相应的极大点和极大值的定义。

下面说明极值点存在的必要条件和充分条件。

定理 1（必要条件） 设 Ω 是 n 维欧氏空间 $1R^n$ 上的某一开集，$f(X)$ 在 Ω 上有一阶连续偏导数，且在点 $X^*\in\Omega$ 处取得局部极值，则必有：

$$\frac{\partial f(X^*)}{\partial x_1}=\frac{\partial f(X^*)}{\partial x_2}=\cdots=\frac{\partial f(X^*)}{\partial x_n}=0 \tag{1.2}$$

或：

$$\nabla f(X^*)=0 \tag{1.3}$$

上式中：

$$\nabla f(X^*)=\left(\frac{\partial f(X^*)}{\partial x_1},\frac{\partial f(X^*)}{\partial x_2},\cdots,\frac{\partial f(X^*)}{\partial x_n}\right)^T \tag{1.4}$$

为函数 $f(x)$ 在点 X^* 处的梯度。

由数学分析知道，$\nabla f(x)$ 的方向为 $f(x)$ 的等值面的法线（在点 X 处）方向，沿这个方向函数值增加最快。

满足式（1.2）或式（1.3）的点称为平稳点或驻点。在区域内部，极值点必为平稳点，但平稳点不一定是极值点。

定理 2（充分条件） 设 Ω 是 n 维欧氏空间 $1R^n$ 上的某一开集，$f(X)$ 在 Ω 上具有二阶连续偏导数，$X^*\in\Omega$，若 $\nabla f(X^*)=0$，且对任何非零向量 $Z\in 1R^n$ 有：

$$Z^TH(X^*)Z>0 \tag{1.5}$$

则 X^* 为 $f(X)$ 的严格局部极小点。此处 $H(X^*)$ 为 $f(X)$ 在点 X^* 处的海赛（Hesse）矩阵：

$$H(X^*)=\begin{pmatrix}\frac{\partial^2 f(X^*)}{\partial x_1^2} & \frac{\partial^2 f(X^*)}{\partial x_1 \partial x_2} \cdots \frac{\partial^2 f(X^*)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(X^*)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(X^*)}{\partial x_2^2} \cdots \frac{\partial^2 f(X^*)}{\partial x_2 \partial x_n} \\ & \cdots\cdots\cdots\cdots \\ \frac{\partial^2 f(X^*)}{\partial x_n \partial x_1} & \frac{\partial^2 f(X^*)}{\partial x_n \partial x_2} \cdots \frac{\partial^2 f(X^*)}{\partial x_n^2}\end{pmatrix} \tag{1.6}$$

为了求解某可微函数的最优解，可令该函数的梯度等于零，由此求得平稳点；然后用充分条件进行判别，求出所要的解。对某些较简单的函数，这样做有时是可行的；但对一般 n 元函数来说，由条件 $\nabla f(x)=0$ 得到的常是一个非线性方程组，解它相当困难。对于不可微函数，当然谈不上使用这样的方法。为此，常直接使用迭代法。

迭代法的基本思想是：为了求函数 $f(X)$ 的最优解，首先给定一个初始估计 $X^{(0)}$，然后按某种规则（即算法）找出比 $X^{(0)}$更好的解 $X^{(1)}$，再按此规则找出比 $X^{(1)}$更好的解 $X^{(2)}$，…，如此即可得到一个解的序列 $\{X^{(K)}\}$。若这个序列有极限 X^*，即：

$$\lim_{K\to\infty}\|X^{(K)}-X^*\|=0 \tag{1.7}$$

则称它收敛于 X^*。

若这算法是有效的，那么它所产生的解的序列将收敛于该问题的最优解。不过，由于计算机只能进行有限次迭代。一般说很难得到准确解，而只能得到近似解。当这一近似解满足所要求的精度时，即可停止迭代。

若由某算法所产生的解的序列 $\{X^{(K)}\}$ 使目标函数值 $f(X^{(K)})$逐步减小，则称这种算法为下降算法。

现假定已经迭代到点 $X^{(K)}$，若从 $X^{(K)}$出发沿任何方向移动都不能使目标函数值下降，则 $X^{(K)}$是一局部极小点，迭代停止。若从 $X^{(K)}$出发至少存在一个方向可使目标函数值有所下降，则可选

定能使目标函数值下降的某方向 $P^{(K)}$，沿这方向迈进适当的一步，得到下一个迭代点 $X^{(K+1)}$，并使 $f(X^{(K+1)})<f(X^{K})$，这相当于在射线：

$$X = X^{(K)} + \lambda P^{(K)} \tag{1.8}$$

上选定新点：

$$X^{(K+1)} = X^{(K)} + \lambda_K P^{(K)} \tag{1.9}$$

其中，$P^{(K)}$称为搜索方向；λ_K 称为步长或步长因子。

下降迭代算法的步骤可总结如下：

(1) 选定某一初始点 $X^{(0)}$，并令 $k=0$。

(2) 确定搜索方向 $P^{(K)}$。

(3) 从 $X^{(K)}$出发，沿方向 $P^{(K)}$求步长 λ_K，以产生下一个迭代点 $X^{(k+1)}$。

(4) 检查得到的新点 $X^{(k+1)}$是否为极小点或近似极小点。若是，则停止迭代。否则，令 $k=k+1$，转回 (2) 继续进行迭代。

在以上步骤中，选取搜索方向 $P^{(K)}$是最关键的一步，各种算法的区分，主要在于确定搜索方向的方法不同。

确定步长 λ_K 可选用不同的方法。最简单的一种是令它等于某一常数（例如令 $\lambda_k=1$），这样做计算简便，但不能保证使目标函数值下降。第二种称为可接受点算法，只要能使目标函数值下降，可任意选取步长 λ_K。第三种方法是沿搜索方向使目标函数值下降最多，即沿射线 $X=X^{(K)}+\lambda P^{(K)}$求目标函数 $f(X)$ 的极小：

$$\lambda_K:\ \min f(X^{(K)}+\lambda P^{(K)}) \tag{1.10}$$

由于这项工作是求以 λ 为变量的一元函数 $f(X^{(K)}+\lambda P^{(K)})$ 的极小点 λ_K，故常称这一过程为（最优）一维搜索或线搜索，这样确定的步长为最佳步长。

一维搜索有个十分重要的性质：在搜索方向上所得最优点处的梯度和该搜索方向正交。

定理3　设目标函数 $f(X)$ 具有一阶连续偏导数，$X^{(K+1)}$按下述规则产生：

$$\left.\begin{aligned}&\lambda_K:\ \min f\left(X^{(K)}+\lambda P^{(K)}\right)\\&X^{(K+1)}=X^{(K)}+\lambda_K P^{(K)}\end{aligned}\right\}\tag{1.11}$$

则有：

$$\nabla f\left(X^{(K+1)}\right)^T P^{(K)}=0\tag{1.12}$$

证明：构造函数 $\phi(\lambda)=f\left(X^{(K)}+\lambda P^K\right)$，则得：

$$\begin{cases}\phi(\lambda_K)=\min\limits_{\lambda}\phi(\lambda)\\X^{(K+1)}=X^{(K)}+\lambda_K P^{(K)}\end{cases}$$

即 λ_K 为 $\phi(\lambda)$ 的极小点。此外：

$$\phi'(\lambda)=\nabla f\left(X^{(K)}+\lambda P^{(K)}\right)^T P$$

由 $\phi'(\lambda)\big|_{\lambda=\lambda K}=0$，可得：

$$\nabla f\left(X^{(K)}+\lambda_K P^{(K)}\right)^T P^{(K)}=\nabla f\left(X^{(K+1)}\right)^T P^{(K)}=0$$

1.2　一维搜索

当用迭代法求函数的极小点时，常常要用到一维搜索，即沿某一已知方向求目标函数的极小点。一维搜索的方法很多，常用的有：(1) 试探法（斐波那契法、0.618法等）；(2) 插值法（抛物线插值法、三次插值法等）；(3) 微积分中的求根法（牛顿法、二分法等）。

1.2.1　牛顿法

由于一维搜索就是求以 λ 为变量的一元函数的极小点，于是，我们的问题就是求解：

$$\min f(x),\quad -\infty<x<+\infty\tag{1.13}$$

或记为：

$$\min f(x),\ x \in R \tag{1.14}$$

假定 $f(x)$ 连续可微，则由极值的必要条件，极小点必为稳定点，这时问题可化为求解方程 $f'(x)=0$，由计算方法中的方程求根的牛顿迭代法，迭代公式为：

$$x^{(K+1)} = x^{(K)} - \frac{f'(x^{(K)})}{f''(x^{(K)})} \tag{1.15}$$

这就是求一维无约束极值的牛顿迭代公式。

牛顿迭代法的计算方法如下：

设已知初始点 x_0，并令 $k=0$。

(1) 计算 $f'(x_K)$，$f''(x_K)$。

(2) 求 $x_{K+1} = x_K - \dfrac{f'(x_K)}{f''(x_K)}$。

(3) 若 $|x_{K+1} - x_K| \leqslant \varepsilon$，则求出最优解的近似解 $x^* = x_{K+1}$，停止计算；否则转 (4)。

(4) 令 $K=K+1$，转 (1)。

注意：二次函数 $\phi(x)$ 仅当 $f''(x_0)>0$ 时才有极小点存在。因此为了保证牛顿法收敛到极小点，应要求 $f''(x_K)>0$，至少对充分大的 K 应如此。

牛顿法的几何解释如图 1－1，曲线 $y=f'(x)$ 在 x_K 处的切线与 x 轴的交点作为下一个点 x_{K+1}。

牛顿法的最大优点是收敛速度快。但是，这个方法也有缺点，要在每个迭代点处计算函数的二阶导数值，这样就增加了每次迭代的工作量。牛顿法的另一个缺点是要求初始点选得比较好，也就是说初始点不能离极小值点太远。

1.2.2 平分法

平分法的收敛速度虽然比牛顿法要慢得多，但每次迭代计算量较少，而且总能收敛到一个局部极小值点。所以它仍是一个受欢迎

的方法。

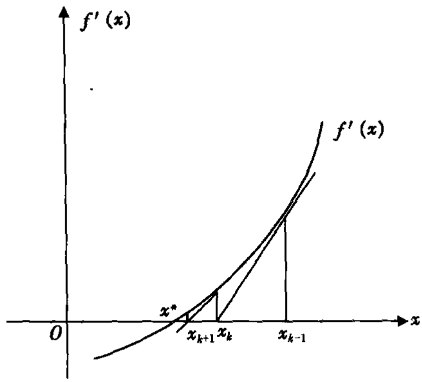

图 1－1

我们知道，在极小点 x^* 处 $f'(x^*)=0$，并且当 $x<x^*$ 时，函数是递减的，即 $f'(x)<0$；而当 $x>x^*$ 时，函数递增，即 $f'(x)>0$。如果我们找到了一个区间 $[a, b]$，且 $f'(a)<0$，$f'(b)>0$，则在 a，b 之间必有 $f(x)$ 的极小点 x^*，并且 $f'(x^*)=0$。为了找到 x^*，我们取 $x_0=\frac{a+b}{2}$，若 $f'(x_0)>0$，则在区间 $[a, x_0]$ 中有极小点，这时以 $[a, x_0]$ 作为新的区间 $[a, b]$；若 $f'(x_0)<0$，则在 $[x_0, b]$ 中有极小点，因此以 $[x_0, b]$ 作为新的区间 $[a, b]$。继续这个过程，逐步将区间 $[a, b]$ 换小，当区间 $[a, b]$ 的长度充分小时，或者当 $f'(x_0)$ 充分小时，即可将 $[a, b]$ 的中点取做极小点的近似点。这时有明显的估计 $\left|x^*-\frac{a+b}{2}\right|<\frac{b-a}{2}$。

至于区间 $[a, b]$ 的确定，一般可采用下述方式。首先取一初始点 x_0，若 $f'(x_0)<0$，则在 x_0 右方取点 $x_1=x_0+\Delta x$（Δx 也是事先给定的一个步长）；若 $f'(x_1)>0$，则令 $a=x_0$，$b=x_1$；若仍有 $f'(x_1)<0$，则取 $x_2=x_1+\Delta x$（或者先将 Δx 增大

一倍，再令 $x_2=x_1+\Delta x$），求 $f'(x_2)$；若 $f'(x_2)>0$，则以 $[x_1, x_2]$ 作为区间 $[a, b]$，否则继续做下去。对 $f'(x_0)>0$ 的情况，也可作类似于 $f'(x_0)<0$ 的情况去讨论，这时应在 x_0 左方取点，取 $\Delta x<0$。

下面给出平分法的计算公式，首先给出找区间 $[a, b]$ 的计算公式。

给定 x_0，Δx 且 $\Delta x>0$

(1) 求 $f'(x_0)$，若 $f'(x_0)=0$，则求出 $x^*=x_0$，停止整个运算；若 $f'(x_0)<0$，做 (2)、(3)；若 $f'(x_0)>0$，进行步骤 (4)、(5)。

(2) $x_1=x_0+\Delta x$（或再做 $\Delta x=\Delta x+\Delta x$）。

(3) 求 $f'(x_1)$，若 $f'(x_1)=0$，则求出 $x^*=x_1$；若 $f'(x_1)>0$，则求出区间 $[a, b]=[x_0, x_1]$；若 $f'(x_1)<0$，则 $x_0=x_1$ 转向 2。

(4) $x_1=x_0-\Delta x$（或再做 $\Delta x=\Delta x+\Delta x$）。

(5) 求 $f'(x_1)$，若 $f'(x_1)=0$，则求出 $x^*=x_1$；若 $f'(x_1)<0$，则求出区间 $[a, b]=[x_1, x_0]$；若 $f'(x_1)>0$，则 $x_0=x_1$ 转向 (4)。

找出了区间 $[a, b]$ 之后，便可用平分法的计算过程。现在我们把其计算方法整理为给定 a，b，ε。

(1) 求 $x=\dfrac{a+b}{2}$，若 $|b-a|\leqslant\varepsilon$，则求出稳定点，即 $x^*=x$；否则做 (2)。

(2) 求 $f'(x)$，若 $f'(x)<0$，则 $a=x$；若 $f'(x)>0$，则 $b=x$，这两种情形均转向 (1)。若 $f'(x)=0$，则求出稳定点 $x=x^*$。

1.2.3 0.618 法（又称黄金分割法）

前面所介绍的牛顿法和平分法均要求计算函数的导数。有时函