

线性约束条件下任意凸函数的神经网络优化模型¹

宋荣方 毕光国*

(南京邮电学院通信工程系 南京 210003)

*(东南大学无线电工程系 南京 210096)

摘要 该文提出了线性约束条件下任意凸函数的神经网络优化模型, 所构造的能量函数的平衡点即为原问题的最优解, 克服了传统的神经网络优化方法所存在的问题, 网络是全局稳定的, 并能收敛到最优点, 计算机仿真结果证明了本文方法的有效性。

关键词 神经网络, 神经优化, 神经计算, 凸规划

中图分类号 TN-052

1 引言

通信领域中, 经常出现线性约束条件下凸函数的实时优化问题, 如自适应天线阵的 MVDR (Minimum Variance Distortionless Response) 波束形成^[1], CDMA 通信中的盲多用户检测^[2], CDMA 系统中的时空综合干扰抑制^[3]以及通信网动态无级路由选择 (Dynamic Non-Hierarchy Routing, DNHR)^[4]等, 由于优化规模都很大, 且必须实时处理, 因此传统的串行计算很难满足要求。而人工神经网络是相互联接的大规模并行处理器, 它能通过高度互联的神经元来实时处理复杂问题, 因此用神经网络进行实时优化具有明显的优越性。

已有许多文献^[5-7]对用神经网络求解约束优化问题作了探讨, 但他们的方法大多是惩罚函数法。根据惩罚定理, 只有当惩罚参数充分大时, 网络的稳态平衡点才能逼近原问题的解。不幸的是, 当惩罚参数很大时, 能量函数的海森矩阵趋向病态, 另外还导致电路实现困难。

本文根据最优化基本理论, 提出了线性约束条件下任意凸函数的神经网络优化一般模型。所构造的能量函数的平衡点为原问题的最优解, 克服了传统的惩罚函数法存在的问题, 同时也克服了普通的拉格朗日乘子法所存在的解的振荡和收敛慢的缺点。本文最后还证明了网络的全局稳定性。

2 能量函数构成

下面首先给出有关的定义和定理, 然后建立线性约束条件下任意凸函数的神经网络优化能量函数。

在欧氏空间 E^n 上定义函数 $f(x)$, $h_i(x)$ 与 $g_j(x)$, 考虑如下问题:

$$\left. \begin{aligned} & \min f(x) \\ & \text{s.t. } h_i(x) = 0, \quad i = 1, 2, \dots, m \\ & \quad g_j(x) \leq 0, \quad j = 1, 2, \dots, p \end{aligned} \right\} \quad (1)$$

在优化理论中有下述定理, 其证明见文献 [8]。

定理 1 一阶必要条件^[8] 如果在 E^n 上定义的函数 $f(x)$, $h_i(x)$ 与 $g_j(x)$ 均有连续的一阶偏导数, x^0 是在约束 $h_i(x) = 0 (i = 1, 2, \dots, m)$, $g_j(x) \leq 0 (j = 1, 2, \dots, p)$ 下 $f(x)$ 的局部 (或

¹ 1999-10-08 收到, 2000-02-14 定稿

国家自然科学基金 (批准号 69882004) 和原邮电部预研基金 (98 部预 4) 资助课题

全局极小点, 且 x^0 是约束条件下的正则点, 令 $l(x) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \sum_{j=1}^p \mu_j g_j(x)$, 则存在 $\lambda_1, \lambda_2, \dots, \lambda_m; \mu_1, \mu_2, \dots, \mu_p$ 使得

$$\nabla l(x^0) = \nabla f(x^0) + \sum_{i=1}^m \lambda_i \nabla h_i(x^0) + \sum_{j=1}^p \mu_j \nabla g_j(x^0) = 0 \quad (2)$$

$$\mu_j g_j(x^0) = 0, \quad j = 1, 2, \dots, p \quad (3)$$

$$\mu_j \geq 0, \quad j = 1, 2, \dots, p \quad (4)$$

以上三式合称为 K-T 条件.

若在问题 (1) 式中, $g_j(x)$ 为 E^n 上的凸函数 ($j = 1, 2, \dots, p$), $h_i(x) = a_i^T x - b_i$ ($i = 1, 2, \dots, m$), 则 $R = \{x | x \in E^n, h_i(x) = 0, i = 1, 2, \dots, m; g_j(x) \leq 0, j = 1, 2, \dots, p\}$ 为凸集, 设 $f(x)$ 为定义在 R 上的凸函数, 此时称

$$\left. \begin{array}{l} \min f(x) \\ \text{s.t. } x \in R \end{array} \right\} \quad (5)$$

为凸规划 [8].

定理 2 凸规划的最优性充分条件 [8] 考虑上述凸规划问题. 设 $f(x), g_j(x)$ ($j = 1, 2, \dots, p$) 有连续的一阶偏导数, 令 $l(x) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \sum_{j=1}^p \mu_j g_j(x)$, 若存在 $x^0, \lambda_1, \lambda_2, \dots, \lambda_m$, 及 $\mu_1, \mu_2, \dots, \mu_p$ 使 $h_i(x^0) = 0, g_j(x^0) \leq 0$ ($i = 1, 2, \dots, m; j = 1, 2, \dots, p$) 且满足 K-T 条件, 则 x^0 是凸规划 (5) 式的最小点.

如果在定理 1 中明确 $h_i(x), g_j(x)$ 为线性函数, 则可取消 “ x^0 是约束条件下的正则点” 这个要求, 从而有下面定理及其证明.

定理 3 一阶必要条件 给出在 E^n 上定义的函数 $f(x), h_i(x)$ 与 $g_j(x)$, $f(x) \in C^1, h_i(x) = a_i^T x - b_i, g_j(x) = c_j^T x - d_j$ ($i = 1, 2, \dots, m; j = 1, 2, \dots, p$), 若 x^0 是 $f(x)$ 在约束条件: $h_i(x) = 0, g_j(x) \leq 0$ ($i = 1, 2, \dots, m; j = 1, 2, \dots, p$) 下的局部 (或全局) 极小点, 令 $l(x) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \sum_{j=1}^p \mu_j g_j(x)$, 则存在 $\lambda_1, \lambda_2, \dots, \lambda_m; \mu_1, \mu_2, \dots, \mu_p$ 使得 K-T 条件成立.

证明 x^0 是在约束条件下 $f(x)$ 的局部 (全局) 极小点, 设 $g_j(x) \leq 0$ ($j \in J$) 为 x^0 的起作用约束, 即

$$h_i(x^0) = 0, \quad g_j(x^0) = 0 \quad i = 1, 2, \dots, m; \quad j \in J \quad (6)$$

设 $R = \{x | h_i(x) = 0, i = 1, 2, \dots, m; g_j(x) \leq 0, j \in J\}$, 则 $x^0 \in R$ 且 x^0 是 R 上 $f(x)$ 的局部 (全局) 极小点.

在 R 上任意取一点 x' , 即

$$h_i(x') = 0, \quad g_j(x') \leq 0 \quad i = 1, 2, \dots, m; \quad j \in J \quad (7)$$

先证明 $x = x^0 + \theta(x' - x^0) \in R$, 其中 $\theta \geq 0$. 设 $z = x' - x^0$, 要证 $x^0 + \theta z \in R$, 即要证 $h_i(x^0 + \theta z) = 0, g_j(x^0 + \theta z) \leq 0$. 因

$$\begin{aligned} h_i(x^0 + \theta z) &= a_i^T(x^0 + \theta z) - b_i = h_i(x^0) + \theta a_i^T z \\ g_j(x^0 + \theta z) &= c_j^T(x^0 + \theta z) - d_j = g_j(x^0) + \theta c_j^T z \\ a_i^T z &= a_i^T(x' - x^0) = (a_i^T x' - b_i) - (a_i^T x^0 - b_i) = h_i(x') - h_i(x^0) \end{aligned}$$

同理 $c_j^T z = g_j(x') - g_j(x^0)$.

由 (6), (7) 式及 $\theta \geq 0$ 得

$$\begin{aligned} a_i^T z &= 0, c_j^T z \leq 0, \quad i = 1, 2, \dots, m; j \in J \\ h_i(x^0 + \theta z) &= \theta a_i^T z = 0, \quad i = 1, 2, \dots, m \\ g_j(x^0 + \theta z) &= \theta c_j^T z \leq 0, \quad j \in J \end{aligned}$$

故 $x^0 + \theta z \in R(\theta \geq 0)$. 由于 x^0 为 R 上 $f(x)$ 的局部 (全局) 极小点, 因此当 θ 充分小时, $f(x^0 + \theta z) \geq f(x^0)$, 于是

$$\lim_{\theta \rightarrow +0} \frac{f(x^0 + \theta z_i) - f(x^0)}{\theta z_i} z_i \geq 0, \text{ 即 } \left. \frac{\partial f(x)}{\partial x_i} \right|_{x=x^0} z_i \geq 0, (i = 1, 2, \dots, n)$$

从而 $(\nabla f(x^0))^T z \geq 0$.

总之, 对满足约束条件, 即 $a_i^T z \geq 0, -a_i^T z \geq 0, -c_j^T z \geq 0, (i = 1, 2, \dots, m; j \in J)$ 的 $z(\theta \geq 0)$ 时, $x^0 + \theta z \in R$, 必有 $(\nabla f(x^0))^T z \geq 0$, 根据 Farkas 引理 [8] 可知 $\nabla f(x^0)$ 必是 $a_i, -a_i, -c_j (i = 1, 2, \dots, m; j \in J)$ 的非负系数线性组合, 即 $\nabla f(x^0) = \sum_{i=1}^m A_i a_i + \sum_{i=1}^m B_i (-a_i) + \sum_{j \in J} \mu_j (-c_j)$, 其中 $A_i \geq 0, B_i \geq 0, \mu_j \geq 0, (i = 1, 2, \dots, m; j \in J)$.

令 $\lambda_i = B_i - A_i$, 且 $j \notin J$ (即 $g_j(x^0) > 0$) 时, 令 $\mu_j = 0$, 则 $\nabla f(x^0) = -\sum_{i=1}^m \lambda_i a_i - \sum_{j=1}^p \mu_j c_j = -\sum_{i=1}^m \lambda_i \nabla h_i(x^0) - \sum_{j=1}^p \mu_j \nabla g_j(x^0)$, 其中 $\mu_j \geq 0$.

由 (6) 式可知, $\mu_j g_j(x^0) = 0, (j = 1, 2, \dots, p)$, 从而 (2), (3), (4) 式成立.

证毕

结合定理 2 和定理 3, 由于线性函数既是凸函数又是凹函数, 因此可得如下定理.

定理 4 线性约束条件下凸规划的最优性充要条件 设凸函数 $f(x) \in C^1$, $h_i(x), (i = 1, 2, \dots, m)$ 与 $g_j(x), (j = 1, 2, \dots, p)$ 都是线性函数, 令 $l(x) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \sum_{j=1}^p \mu_j g_j(x)$, 则 x^0 是凸规划 (5) 式的最小点的充要条件为存在 $\lambda_1, \lambda_2, \dots, \lambda_m; \mu_1, \mu_2, \dots, \mu_p$ 使 x^0 满足约束条件且 K-T 条件成立.

定理 4 是我们导出能量函数的基础, 我们应导出这样的能量函数, 使其平衡点恰好满足定理 4 中的充要条件, 从而获得最优解.

我们构造如下能量函数:

$$E(x, y, \lambda, \mu) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \sum_{j=1}^p \mu_j (g_j(x) + y_j^2) \quad (8)$$

其中 y_j 为辅助变量, $\lambda_i, \mu_j (\mu_j \geq 0, j = 1, 2, \dots, p)$ 为乘子变量. 该能量函数的平衡点由下列方程组确定.

$$\left. \begin{aligned}
 \nabla_{x_k} E(x^0, y^0, \lambda^0, \mu^0) &= \nabla_{x_k} f(x^0) + \sum_{i=1}^m \lambda_i^0 \nabla_{x_k} h_i(x^0) \\
 &+ \sum_{j=1}^p \mu_j^0 \nabla_{x_k} g_j(x^0) = 0, \quad k = 1, 2, \dots, n \\
 \nabla_{\lambda_i} E(x^0, y^0, \lambda^0, \mu^0) &= h_i(x^0) = 0, \quad i = 1, 2, \dots, m \\
 \nabla_{\mu_j} E(x^0, y^0, \lambda^0, \mu^0) &= g_j(x^0) + y_j^{0^2} = 0, \quad j = 1, 2, \dots, p \\
 \nabla_{y_j} E(x^0, y^0, \lambda^0, \mu^0) &= 2\mu_j^0 y_j^0 = 0, \quad j = 1, 2, \dots, p \\
 \mu_j^0 &\geq 0, \quad j = 1, 2, \dots, p
 \end{aligned} \right\} \quad (9)$$

比较以上方程组和定理 4 中的充要条件, 因为

$$g_j(x^0) + y_j^{0^2} = 0 \Rightarrow \left\{ \begin{array}{l} g_j(x^0) \leq 0; \quad \mu_j^0 [g_j(x^0) + y_j^{0^2}] = 0 \\ \mu_j^0 [g_j(x^0) + y_j^{0^2}] = 0; \quad \mu_j^0 y_j^0 = 0 \end{array} \right\} \Rightarrow \mu_j^0 g_j(x^0) = 0;$$

以及 $g_j(x^0) \leq 0 \Rightarrow \exists y_j^0$, 使 $g_j(x^0) + y_j^{0^2} = 0 \Rightarrow \left. \begin{array}{l} \mu_j^0 [g_j(x^0) + y_j^{0^2}] = 0 \\ \mu_j^0 g_j(x^0) = 0 \end{array} \right\} \Rightarrow \mu_j^0 y_j^0 = 0$, 所以定理 4

中的充要条件等价于能量函数 $E(x, y, \lambda, \mu)$ 的平衡点条件, 且关于该平衡点, 有下述定理成立.

定理 5 能量函数 (8) 式满足如下不等式: $E(x^0, y^0, \lambda, \mu) \leq E(x^0, y^0, \lambda^0, \mu^0) \leq E(x, y, \lambda^0, \mu^0)$, 即该函数关于 x, y 存在最小点 (x^0, y^0) , 关于 λ, μ 存在最大点 (λ^0, μ^0) .

证明 因为 $\mu_j \geq 0, h_i(x), g_j(x)$ 均为线性函数, y_j^2 为凸函数, 所以 $E(x, y, \lambda, \mu)$ 为关于 x, y 的凸函数, 因此, 对于满足 (9) 式的平衡点, 必满足:

$$E(x^0, y^0, \lambda^0, \mu^0) = \min_{x, y} E(x, y, \lambda^0, \mu^0) \quad (\text{全局最小})$$

又 $h_i(x^0) = 0, (i = 1, 2, \dots, m); g_j(x^0) + y_j^{0^2} = 0, (j = 1, 2, \dots, p)$.

根据鞍点的充要条件^[8]可知, $E(x^0, y^0, \lambda^0, \mu^0)$ 为 $E(x, y, \lambda, \mu)$ 的鞍点, 根据鞍点定义, 上述不等式成立, 从而定理得证. 证毕

为了加快收敛速度, 减轻解的振荡, 可在能量函数中加上惩罚函数, 但不要求惩罚参数很大, 从而避免了单纯使用惩罚函数时所面临的问题. 加上惩罚函数项的能量函数为

$$E(x, y, \lambda, \mu) = f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \sum_{j=1}^p \mu_j (g_j(x) + y_j^2) + \frac{\alpha}{2} \left[\sum_{i=1}^m h_i^2(x) + \sum_{j=1}^p (g_j(x) + y_j^2)^2 \right] \quad (10)$$

其中 $\alpha > 0$ 为惩罚参数, $\mu_j \geq 0, j = 1, 2, \dots, p$.

3 神经网络模型

根据前面的讨论, 线性约束条件下任意凸函数的优化问题等价成如下问题:

$$\max_{\lambda, \mu} \min_{x, y} E(x, y, \lambda, \mu), \quad \mu_j \geq 0, j = 1, 2, \dots, p$$

对 (10) 式采用梯度策略, 即可获得神经网络的状态方程 (11) 式, 但须注意, 对于 x , y 采用梯度下降法, 而对于 λ , μ 采用梯度上升法.

$$\left. \begin{aligned} \frac{dx_k}{dt} &= -\rho \frac{\partial E}{\partial x_k} = -\rho \left[\frac{\partial f(x)}{\partial x_k} + \sum_{i=1}^m \lambda_i \frac{\partial h_i(x)}{\partial x_k} + \sum_{j=1}^p \mu_j \frac{\partial g_j(x)}{\partial x_k} + \alpha \sum_{i=1}^m h_i(x) \frac{\partial h_i(x)}{\partial x_k} \right. \\ &\quad \left. + \alpha \sum_{j=1}^p (g_j(x) + y_j^2) \frac{\partial g_j(x)}{\partial x_k} \right], \quad k = 1, 2, \dots, n \\ \frac{dy_j}{dt} &= -\rho \frac{\partial E}{\partial y_j} = -\rho [2\mu_j y_j + 2\alpha (g_j(x) + y_j^2) y_j], \quad j = 1, 2, \dots, p \\ \frac{d\lambda_i}{dt} &= \rho \frac{\partial E}{\partial \lambda_i} = \rho h_i(x), \quad j = 1, 2, \dots, m \\ \frac{d\mu_j}{dt} &= \rho \frac{\partial E}{\partial \mu_j} = \rho [g_j(x) + y_j^2], \quad j = 1, 2, \dots, p \\ \mu_j &\geq 0; \quad j = 1, 2, \dots, p; \quad \rho > 0 \end{aligned} \right\} \quad (11)$$

以上状态方程很容易实现, 每个神经元为一有源求和积分网络, 为了限制 $\mu_j \geq 0 (j = 1, 2, \dots, p)$, 只要在相应神经元输出端与地之间加上钳位二极管即可.

关于网络的稳定性, 有下述定理.

定理 6 由 (11) 式表示的网络是全局 Lyapunov 稳定的, 并收敛到最优点.

证明 为了简洁, 我们全部采用矢量形式表示, 并设 $\rho = 1$.

$$\left. \begin{aligned} \frac{dx}{dt} &= \left[\frac{dx_1}{dt}, \frac{dx_2}{dt}, \dots, \frac{dx_n}{dt} \right]^T = - \left[\frac{\partial E}{\partial x_1}, \frac{\partial E}{\partial x_2}, \dots, \frac{\partial E}{\partial x_n} \right]^T = -\nabla_x E \\ \frac{dy}{dt} &= \left[\frac{dy_1}{dt}, \frac{dy_2}{dt}, \dots, \frac{dy_p}{dt} \right]^T = - \left[\frac{\partial E}{\partial y_1}, \frac{\partial E}{\partial y_2}, \dots, \frac{\partial E}{\partial y_p} \right]^T = -\nabla_y E \\ \frac{d\lambda}{dt} &= \left[\frac{d\lambda_1}{dt}, \frac{d\lambda_2}{dt}, \dots, \frac{d\lambda_m}{dt} \right]^T = [h_1(x), h_2(x), \dots, h_m(x)]^T = h(x) \\ \frac{d\mu}{dt} &= \left[\frac{d\mu_1}{dt}, \frac{d\mu_2}{dt}, \dots, \frac{d\mu_p}{dt} \right]^T = [g_1(x) + y_1^2, g_2(x) + y_2^2, \dots, g_p(x) + y_p^2]^T = h(x, y) \end{aligned} \right\} \quad (12)$$

我们把 Lyapunov 函数定义成

$$L(x, y, \lambda, \mu) = \frac{1}{2} \|\nabla_x E\|^2 + \frac{1}{2} \|\nabla_y E\|^2 + \frac{1}{2} \|h(x)\|^2 + \frac{1}{2} \|h(x, y)\|^2 \quad (13)$$

则

$$\frac{dL}{dt} = \left[\frac{\partial L}{\partial x} \right]^T \frac{dx}{dt} + \left[\frac{\partial L}{\partial y} \right]^T \frac{dy}{dt} + \left[\frac{\partial L}{\partial \lambda} \right]^T \frac{d\lambda}{dt} + \left[\frac{\partial L}{\partial \mu} \right]^T \frac{d\mu}{dt} \quad (14)$$

根据矢量求导法则对 (13) 式求导并整理得

$$\left. \begin{aligned} \frac{\partial L}{\partial x} &= \nabla_{xx}^2 E \nabla_x E + \left[\frac{\partial h(x)}{\partial x} \right]^T h(x) + \left[\frac{\partial h(x,y)}{\partial x} \right]^T h(x,y) \\ \frac{\partial L}{\partial y} &= \nabla_{yy}^2 E \nabla_y E + \left[\frac{\partial h(x,y)}{\partial y} \right]^T h(x,y) \\ \frac{\partial L}{\partial \lambda} &= \frac{\partial h(x)}{\partial x} \nabla_x E \\ \frac{\partial L}{\partial \mu} &= \frac{\partial h(x,y)}{\partial x} \nabla_x E + \frac{\partial h(x,y)}{\partial y} \nabla_y E \end{aligned} \right\} \quad (15)$$

把 (12), (15) 式代入 (14) 式并整理得

$$\frac{dL}{dt} = -\nabla_x E^T \nabla_{xx}^2 E \nabla_x E - \nabla_y E^T \nabla_{yy}^2 E \nabla_y E$$

因为 E 是 x, y 的凸函数, 所以 $\nabla_{xx}^2 E, \nabla_{yy}^2 E$ 均为非负定矩阵, 从而 $\frac{dL}{dt} \leq 0$; 又因为 $L \geq 0$, 故 $\lim_{t \rightarrow \infty} \frac{dL}{dt} = 0$, 它意味着 $\nabla_x E = 0, \nabla_y E = 0, h(x) = 0, h(x, y) = 0$. 证毕

最后需说明的是, 由于在能量函数 (8) 和 (10) 式中引入了辅助变量, 所以神经元个数增加了, 我们可消去辅助变量而获得新的网络模型, 但将引入开关函数, 从而网络实现时要用到比较器和开关, 有兴趣的读者可参阅文献 [9].

4 计算机仿真

线性规划和凸二次规划是线性约束条件下凸规划的典型实例, 下面我们以二次规划为例验证本文方法的有效性.

例 1

$$\begin{aligned} \min f(x) &= x_1^2 + x_2^2 + x_3^2 + x_4^2 - 2x_1 - x_4 \\ \text{s.t. } &2x_1 + x_2 + x_3 + 4x_4 = 7; \quad 0.2x_1 + 2x_2 + 2x_3 = 6 \end{aligned}$$

其理论解为 $x_1=1.1275, x_2=1.4436, x_3=1.4436, x_4=0.4644$.

图 1(a) 为 $\rho = 10^6, \alpha = 1$, 时长 $8\mu\text{s}$, 初始值 $x_1=0, x_2=0, x_3=0, x_4=0, \lambda_1=0, \lambda_2=0$ 时的仿真曲线, 获得的最优解为 $x_1=1.1275, x_2=1.4434, x_3=1.4434, x_4=0.4646, \lambda_1=0.0175, \lambda_2=-1.4515$.

图 1(b) 为 $\rho = 10^6, \alpha = 0$, 即不含惩罚函数项, 时长 $8\mu\text{s}$, 初始值同上时的仿真曲线, 获得的最优解为 $x_1=1.1273, x_2=1.4431, x_3=1.4431, x_4=0.4642, \lambda_1=0.0178, \lambda_2=-1.4522$.

图 1(c) 为 $\rho = 10^6, \alpha = 1$, 且单纯使用惩罚函数项, 时长 $8\mu\text{s}$, 初始值同上时的仿真曲线, 获得的解为 $x_1=1.1552, x_2=1.1579, x_3=1.1579, x_4=0.5830$.

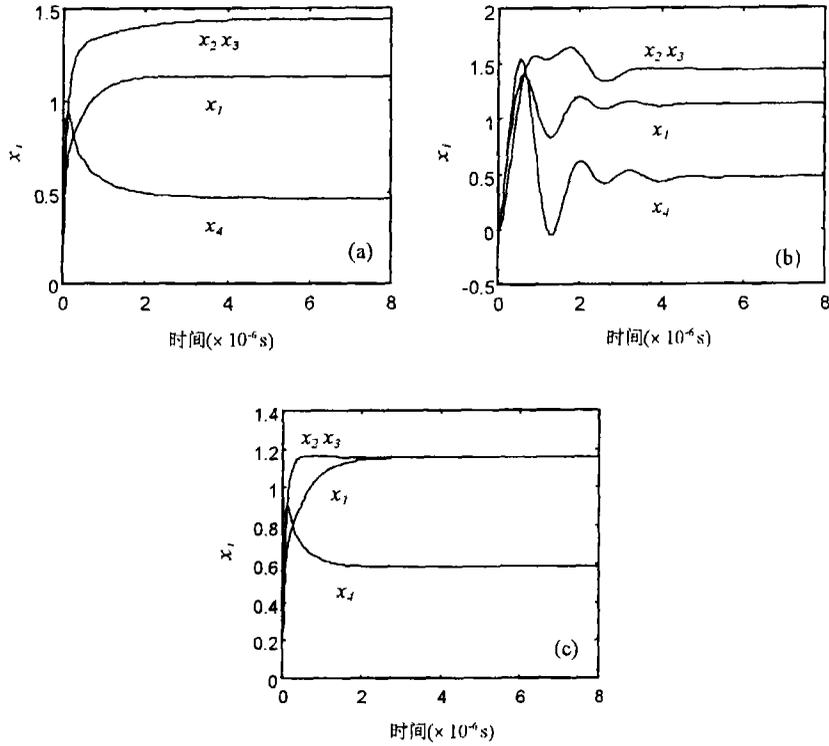


图 1 例 1 的仿真结果

例 2

$$\min f(x) = x_1^2 - x_1x_2 + 2x_2^2 - x_1 - x_2$$

$$\text{s.t. } x_1 + x_2 - 4 = 0; \quad x_1 - x_2 - 3 \geq 0$$

其理论解为 $x_1=3.5, x_2=0.5$.

设 $\rho = 10^6, \alpha = 1$, 时长 $20\mu\text{s}$, 图 2(a) 为初始值 $x_1=0, x_2=0, \lambda=0, \mu=0, y=0$ 时的仿真曲线, 获得的最优解为 $x_1=3.4996, x_2=0.5011, \lambda=-1.5017, \mu=3.9960, y=0$. 图 2(b) 为选择不同初始值时 x_1-x_2 的轨迹图.

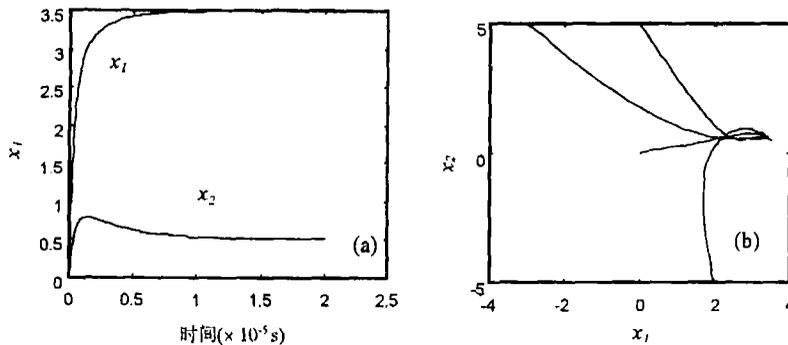


图 2 例 2 的仿真结果

通过对大量实例的仿真, 本文方法均达到了理想结果, 初始值可在可行域内外任意选取, 另外我们也看到, 增加惩罚函数项能减轻解的振荡, 加快收敛速度。而单纯使用惩罚函数法时, 获得的解误差较大。

参 考 文 献

- [1] S. Haykin, Adaptive Filter Theory, 3rd ed. NJ: Prentice-Hall, 1996, Chapter 5.
- [2] J. B. Schodorf, D. B. Williams, A constrained optimization approach to multiuser detection, IEEE Trans. on SP, 1997, SP-45(1), 258-262.
- [3] 宋荣方, 毕光国, CDMA 系统中时空综合干扰抑制的约束优化模型及其神经网络实现, 电子学报, 2001, 29(4), 475-478.
- [4] 阙超, 基于神经网络的电信网 DNHR 算法及实现 DNHR 的有关协议, [硕士论文], 南京, 南京邮电学院, 1991.
- [5] D. D. Tank, J. J. Hopfield, Simple neural optimization networks: An A/D converter, signal decision and a linear programming, IEEE Trans. on CAS, 1986, 33(5), 533-541.
- [6] M. P. Kenney, L. O. Chua, Neural network for nonlinear programming, IEEE Trans. on CAS, 1988, 35(5), 554-562.
- [7] A. Rodriguez, *et al.*, Nonlinear switched-capacitor neural network for optimization problems, IEEE Trans. on CAS, 1990, 37(3), 384-397.
- [8] 陈开明, 非线性规划, 上海, 复旦大学出版社, 1991, 第 2 章.
- [9] 宋荣方, 毕光国, 线性约束凸规划神经网络新模型, 信号处理, 2001, 17(2), 104-109.

A NEURAL NETWORK MODEL FOR THE OPTIMIZATION OF ARBITRARY CONVEX FUNCTIONS WITH LINEAR CONSTRAINTS

Song Rongfang Bi Guangguo*

(Dept. of Telecommun., Nanjing Inst. of Posts and Telecommunications, Nanjing 210003, China)

*(Dept. of Radio Engineering, Southeast University, Nanjing 210096, China)

Abstract This article presents a neural network model for the optimization of arbitrary convex functions with linear constraints. The equilibrium point of the energy function constructed is the optimal solution of the original problem. The problems, which would arise in conventional neural network optimization methods, are overcome. The neural model is globally stable and can converge to the optimal point. The computer simulation results verify the effectiveness of the method.

Key words Neural network, Neural optimization, Neural computing, Convex programming

宋荣方: 男, 1964 年生, 副教授, 博士, 主要从事智能信号处理在通信中的应用和空分 / 码分多址通信的研究。

毕光国: 男, 1939 年生, 教授, 博士生导师, 主要从事码分多址和宽带数字移动通信、现代通信中的信号处理和智能信息处理以及宽带多媒体通信的研究。