

文章编号:1000-6788(2003)09-0098-07

求解有界约束二次规划问题神经网络模型的收敛性分析

李有梅^{1,2}, 彭济根¹, 徐宗本¹

(1. 西安交通大学理学院信息与系统科学研究所, 陕西 西安 710049; 2. 山西大学计算机科学系, 山东 太原 030006)

摘要: 通过深化 Lasalle 不变原理, 建立了判别一般动力系统全局收敛性的一个准则. 应用这一准则, 详尽研究了一个求解有界约束二次规划问题神经网络的全局收敛性, 给出了当目标函数为一类非凸函数时的全局收敛性条件. 特别地利用常微分方程理论, 证明了该网络对任意凸函数全局收敛性, 所获结果深化和推广了现有文献相关结论的相应结论. 这些新的结论都表明了该神经网络在求解有界约束二次规划问题时的有效性. 数值模拟与理论分析结果一致.

关键词: 神经网络; 二次优化; 全局收敛; 平衡点

中图分类号: TP18

文献标识码: A

Convergence Analysis of a Neural Network for Quadratic Optimization with Bound Constraints

LI You-mei^{1,2}, PENG Ji-gen¹, XU Zong-ben¹

(1. Institute for Information and System Science, School of Science, Xi'an Jiaotong University, Xi'an 710049, China; 2. Department of Computer Science, Shanxi University, Taiyuan, 030006, China)

Abstract: In this paper, we present a general principle by sharpening lasalle invariance principle to judge the convergence of a dynamic system. Based on this, The global convergence of a neural network for quadratic optimization with bound constraints is studied in detail, and some new conditions are obtained on which the neural network can be guaranteed to be globally convergent for a non-convex objective function. Specially using ordinary differential equation theory, the global convergence of the neural network is proved when objective function is convex, this conclusion generalizes and deepens the corresponding results obtained in current literaturess. All these new results show the validity of the network in solving quadratic optimization with bound constraints. Finally, two numerical examples are given to demonstrate the correctness of the theoretical results.

Key words: neural network; quadratic optimization; global convergence; equilibrium point

1 引言

考虑有界约束二次规划问题

$$\min\{E(x) = \frac{1}{2}x^T Qx - x^T q; x \in \Omega \subset R^N\}$$

这里约束集 $\Omega = \{(x_1, x_2, \dots, x_N)^T, x_i \in [a_i, b_i]\}$. 这类问题广泛出现在科学与工程领域中, 例如, 机器人运动实时控制问题, 就可以描述为对时变的如上二次规划问题的求解. 神经网络的快速求解能力及其硬件可实现性使之成为求解此类问题的有效方法^[1,2]. 如何有效地求解该类问题一直受到人们的普遍关注. 近年来, 对求解该类问题的神经网络方法研究异常热烈, 且已经取得一系列主要成果. 例如, [3-5]提出了基于罚函数法和拉格朗日乘子法的神经网络模型, 这些模型虽然具有处理问题广泛的优点, 但不保证所生成

收稿日期: 2001-04-26

资助项目: 国家自然科学基金(69975016)

作者简介: 李有梅(1965-), 女, 汉族, 山西大同人, 副教授, 博士研究生, 主要研究方向是神经网络和演化计算; 徐宗本(1955-), 男, 汉族, 博士生导师, 主要研究方向是神经网络和演化计算

解的可行性. 文献[6—8]相继提出并研究下述模型:

$$\tau \frac{dx}{dt} = -x + P_{\Omega}(x - \alpha Qx + \alpha q) \quad (1)$$

其中, $\tau = \text{diag}(\tau_1, \tau_2, \dots, \tau_n)$, $\alpha = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_n)$ 为两个参数正对角矩阵, Q 是 N 阶对称矩阵. 模型中激活函数 $P_{\Omega}(u) = (g_1(u_1), g_2(u_2), \dots, g_n(u_n))^T$ 定义为:

$$g_i(x_i) = \begin{cases} a_i, & \text{if } u_i < a_i \\ x_i, & \text{if } a_i \leq u_i \leq b_i, \\ b_i, & \text{if } b_i < u_i. \end{cases}$$

$u = (u_1, u_2, \dots, u_N)^T$, u_i 是神经元 i 的内部状态变量, $x_i = g_i(u_i)$ 是神经元 i 的输出状态变量. 容易看到函数 $P_{\Omega}(u)$ 是 $R^N \rightarrow \Omega$ 上的投影算子, 而且由于 Ω 是凸集, 向量 $-x + P_{\Omega}(x - \alpha Qx + \alpha q)$ 是所给二次规划问题的一个可行搜索方向. 所以模型(1)可理解作可行下降优化方法的神经网络实现, 故它必然保证模型求得问题的可行解. 另外, 与基于梯度下降法和罚函数法所设计的模型相比, 模型(1)在硬件实现上也具有显著优点.

本文将主要研究神经网络模型(1)的全局收敛性质. 熟知, 神经网络的全局收敛性是神经网络可用于求解优化问题的前提, 它保证从任何初始点起始, 网络可正确输出问题的一个可行解. 在文献[6, 7]中, 作者证明了在 Q 为对称正定情形模型(1)具有全局指数收敛性, 以及在半正定情形下的拟收敛性. 但在 Q 半正定或不定情形, 对网络的全局收敛性没有获得任何有用的结论. 注意到, Q 正定对应于目标函数 $E(x)$ 是一致凸函数, 而 Q 半正定推出 $E(x)$ 是一般的凸函数, 所以澄清在 Q 半正定情形网络(1)的全局收敛性有重要意义与应用价值.

本文将通过对 Lasella 不变原理的深化, 建立在一般情形下模型(1)全局收敛的充分性条件, 并特别证明在半正定情形模型(1)的全局收敛性. 所获结果将不仅从理论上推广[6—8]的工作, 而且将揭示模型(1)作为求解有界约束二次规划问题算法的可行性与有效性. 本文安排如下: 第二节我们给出一个收敛判别准则; 在第三节中具体给出模型(1)的收敛性分析; 第四节给出模型(1)在正半定情形下的收敛性证明; 第五节通过模拟实例解释网络的性能, 最后在第六节总结全文.

2 一个一般的收敛性判别准则

为了后面准确地表述和理解, 我们先确认相关的概念与定义. 考虑一般的自治系统:

$$\frac{dx}{dt} = F(x), \quad x \in \Omega \subset R^N \quad (2.1)$$

我们知道, 任意给定初始点 x_0 , 当 $F(x)$ 是局部 Lipschitz 函数时, 系统(2.1)存在唯一的解 $x(t, x_0)$. 解 $x(t, x_0)$ 也常被称作是一条通过点 x_0 轨线, 记为 $\Gamma(x_0)$. 如果存在点 x^* 满足 $F(x^*) = 0$, 称 x^* 是系统(2.1)的平衡点, (2.1)的全体平衡点的集合记为 $F^{-1}(0)$. 对于 Ω 的一个子集 D , 如果 $x_0 \in D$ 保证 $\Gamma(x_0) \subset D$, 则称 D 为系统的一个不变集. 点 \bar{x} 称为是 $\Gamma(x_0)$ 的一个 ω 极限点, 是指存在一个子序列 $\{t_i\}$ 使得 $\bar{x} = \lim_{i \rightarrow \infty} x(t_i, x_0)$. 所有的 ω 极限点组成(2.1)的 ω 极限集, 记为 $\omega(\Gamma(x_0))$, 已熟知, ω 极限集是系统的不变集. 涉及优化问题求解时, 我们主要关心的是相应系统的全局收敛性. 下面给出这一概念的一个准确定义:

定义 1 若对任意的 $x_0 \in \Omega$, 轨线 $x(t, x_0)$ 当 $t \rightarrow \infty$ 时都有极限 $x^*(x_0)$, 则称系统(2.1)关于 Ω 是全局收敛的; 若对(2.1)的某个不变子集 M , $x(t, x_0)$ 与集合 M 的距离当 $t \rightarrow \infty$ 时趋于零, 则称(2.1)为拟收敛的.

根据上述定义, 一个全局收敛的神经网络必然保证该网络对任何初值收敛(通常收敛极限点正给出所求解问题的解), 但对不同的初值, 其收敛极限可以不同. 另外, 一个拟收敛的网络可以本身是振荡的, 因而可以不收敛. 应用中, 人们自然希望有一个准则能判定一个系统的收敛性. 下述 Lasalle 不变原理提供了这种判定的理论基础.

Lasalle 不变原理 设 D 是一个有界闭集且从 D 内出发的(2.1)的解 $x(t, x_0) \in D$. 又设存在一个函数 $E(x): D \rightarrow R$, 它具有一阶连续偏导数, 且使得:

$$\dot{E}(x(t)) = \frac{dE(x(t), x_0)}{dt} \leq 0$$

记 $E_0 = \{x, |\dot{E}(x) = 0, x \in D\}$, 则存在常数 c 使得 $\omega(\Gamma(x_0)) \subseteq M \cap E^{-1}(c)$, 且 $x(t, x_0)$ 拟收敛到 M . 这里 M 是包含在 E_0 中的最大不变子集, $E^{-1}(c) = \{x; E(x) = c\}$.

函数 $E(x)$ 称为系统 (2.1) 的能量函数, 它沿着轨线 $x(t, x_0)$ 的导数非正, 函数值是不增的. 如果 $\dot{E}(x(t))|_{x^*} = 0$, 当且仅当 $x^* \in F^{-1}(0)$, 称 $E(x)$ 为系统 (2.1) 的严格能量函数. Lasalle 不变原理虽然提供了系统 (2.1) 的 ω 极限集的定位信息, 断定了系统 (2.1) 的拟收敛性, 但对轨线的收敛性却没有给出任何具体结论. 即使在拟收敛的情形下, 轨线仍有可能是震荡的. 以下我们将深化 Lasalle 不变原理, 以得到系统 (2.1) 的收敛性结论. 以下我们用 $\langle \cdot, \cdot \rangle$ 表示通常的内积, 而用 $\nabla E(x)$ 表示 $E(x)$ 在 x 处的梯度.

定理 1 若自治系统 (2.1) 对任意初始点 x_0 存在唯一有界解 $x(t, x_0)$, 并满足条件:

- 1) 系统有一个严格的能量函数;
- 2) 集合 $F^{-1}(0) \cap \omega(\Gamma(x_0))$ 的元素 $|F^{-1}(0) \cap \omega(\Gamma(x_0))|$ 为至多可数无穷;

则 $x(t, x_0)$ 必收敛到 (2.1) 的一个平衡点.

证明 由条件 (1) 和 Lasalle 不变原理, 我们知道, $\omega(\Gamma(x_0)) \subseteq F^{-1}(0)$. 因此, 我们只需证明 $\omega(\Gamma(x_0))$ 是单点集即可. 采用反证法. 记 $x(t, x_0) = (x_1(t), x_2(t), \dots, x_N(t))^T$. 假设至少存在两个元素 $x^{1*}, x^{2*} \in \omega(\Gamma(x_0))$ 且它们至少有一个分量不同. 记 $x^{1*} = (x_1^{1*}, x_2^{1*}, \dots, x_N^{1*})^T$, $x^{2*} = (x_1^{2*}, x_2^{2*}, \dots, x_N^{2*})^T$. 且不妨设它们的第一个分量不相等, $x_1^{1*} \neq x_1^{2*}$. 由极限点的定义可知, 存在两个子列 $\{t_i\}, \{s_i\}$, 使得 $\lim_{i \rightarrow \infty} x(t_i, x_0) = x^{1*}, \lim_{i \rightarrow \infty} x(s_i, x_0) = x^{2*}, x_1(t_i) \rightarrow x_1^{1*}$, 且 $x_1(s_i) \rightarrow x_1^{2*}$ (当 $i \rightarrow \infty$). 对任意的 $\lambda \in (0, 1)$, 令 $x_1^{\lambda*} = \lambda x_1^{1*} + (1 - \lambda)x_1^{2*}$. 由于 $x(t, x_0)$ 是连续的, 根据介值定理, 存在子列 $\gamma_i^1 \in (\min(t_i, s_i), \max(t_i, s_i))$, 当 γ_i^1 足够大时, 有 $x_1(\gamma_i^1) = x_1^{\lambda*}$, 当然有 $\lim_{i \rightarrow \infty} x(\gamma_i^1, x_0) = x_1^{\lambda*}$. 对于 $\{\gamma_i^1\}$, 由于 $x(t, x_0)$ 是有界的, 所以 $x_2(t)$ 有界, 当然 $x_2(\gamma_i^1)$ 有界, 故必有 γ_i^1 的子列 γ_i^2 , 使得 $x_2(\gamma_i^2) \rightarrow a, (i \rightarrow \infty)$. 令 $x_2^{\lambda*} = a$. 同理, 对 $x_3(t), x_4(t), \dots, x_N(t)$ 逐步抽取收敛子列, 并记它们的极限分别为 $x_3^{\lambda*}, x_4^{\lambda*}, \dots, x_N^{\lambda*}$, 最终我们得到一子列 $\{\gamma_i^N\}$ 且 $x(\gamma_i^N, x_0) \rightarrow x^{\lambda*} = (x_1^{\lambda*}, x_2^{\lambda*}, \dots, x_N^{\lambda*})$. 这表明, 对任何 $\lambda \in (0, 1)$, $x^{\lambda*}$ 都是一个极限点. 由于 $x_1^{1*} \neq x_1^{2*}$, 且 λ 是任意的, 故 $x^{\lambda*}$ 有无穷多个. 这推出 $\omega(\Gamma(x_0))$ 包含有不可数个元素, 与条件 2) 矛盾. (证毕)

定理 1 是一个一般的全局收敛性判别准则. 针对具体的模型 (1), 我们将在以下几节应用定理 1 得出它的若干具体的收敛性结论.

3 模型 (1) 的收敛性分析: 不定情形

本节我们应用上述定理 1 证明在 Q 不定情形下模型 (1) 的若干收敛性定理. 首先我们注意到, 投影算子 $P_\Omega(x)$ 具有下述的性质:

$$\| (P_\Omega(x) - P_\Omega(y)) \|^2 \leq \langle (P_\Omega(x) - P_\Omega(y)), (x - y) \rangle \leq \| x - y \|^2$$

该性质推出 $P_\Omega(x)$ 是 Lipschitz 连续的, 从而模型 (1) 对任意 x_0 存在唯一解 $x(t, x_0)$.

我们需要下述几个引理:

引理 1 Ω 是模型 (1) 的不变集, 即对于任何 $x_0 \in \Omega$, 轨线 $x(t, x_0) \in \Omega$.

证明见 [6].

引理 2 若矩阵 Q 满足下述条件 C: 则模型 (1) 在 Ω 内有且仅有有限个平衡点.

C: Q 的任意阶主子矩阵非奇异, 或下述任一子条件 $C_1 - C_4$ 满足.

$$C_1: \| I - Q \| < 1;$$

$$C_2: |Q_{ii}| > \max \left\{ \sum_{j=1}^N \alpha_i / \alpha_j * |Q_{ij}|, \sum_{j=1}^N \alpha_i / \alpha_j * |Q_{ji}| \right\};$$

$$C_3: Q \text{ 或 } -Q \text{ 是非奇异的 M 矩阵};$$

$$C_4: Q \text{ 或 } -Q \text{ 是正定矩阵};$$

证明 由于 $P_\Omega(x - aQx + aq)$ 显然是映 Ω 入自身的连续映射,且 Ω 是有界闭凸集,故依熟知的 Brouwer 不动点定理,它在 Ω 内必存在不动点,从而系统(1)有平衡点存在. 下证这样的平衡点只有有限个. 我们首先给出 Ω 的 k -维面的定义: 给定指标集合 $I_k = \{i_1, i_2, \dots, i_k\}$, Ω 的一个 k -维面定义为:

$$\Omega_{I_k} = \{(x_1, x_2, \dots, x_N) \mid \forall i \in I_k, x_i = a_i \text{ or } b_i; \forall i \in N \setminus I_k, a_i < x_i < b_i\}$$

亦即 Ω_{I_k} 中恰有 k 个分量固定为端点值. 显然 Ω 的 k -维面的个数是有限的,且 k 可取值为 $0, 1, \dots, N-1$.

为了证明系统(1)有有限个平衡点,我们以下证明在每个 k -维面上系统(1)最多只有一个平衡点. 假设不然,在某一个 k -维面 Ω_{I_k} 上系统(1)有两个平衡点 $x_1^* = (x_{11}^*, x_{12}^*, \dots, x_{1N}^*)$ 和 $x_2^* = (x_{21}^*, x_{22}^*, \dots, x_{2N}^*)$. 于是由平衡点的定义,我们有: 对于任何 $i \in I_k, x_{1i}^* = x_{2i}^*$; 而对于任何 $i \in N \setminus I_k$,

$$x_{1i}^* = P_\Omega(x_{1i}^* - (aQ)_i x_{1i}^* - a_i q_i)$$

$$x_{2i}^* = P_\Omega(x_{2i}^* - (aQ)_i x_{2i}^* - a_i q_i)$$

且必存在 i_0 使得 $x_{1i_0}^* \neq x_{2i_0}^*$. 注意到 $a_i < x_{1i}^*, x_{2i}^* < b_i$ 以及投影算子的性质,我们于是得出

$$x_{1i}^* = (x_{1i}^* - (aQ)_i x_{1i}^* - a_i q_i)$$

$$x_{2i}^* = (x_{2i}^* - (aQ)_i x_{2i}^* - a_i q_i)$$

这里 $(aQ)_i$ 是矩阵 aQ 的第 i 行. 令 $J_k = N \setminus I_k$, 则上述方程可统一写成矩阵表达式:

$$(aQ)_{J_k \times J_k} (x_1^* - x_2^*)_{J_k} = 0$$

其中 $(aQ)_{J_k \times J_k}$ 表示 aQ 的 J_k 阶主子式,且 $(x_1^* - x_2^*)_{J_k}$ 是由 J_k 中指标相应的 $(x_1^* - x_2^*)$ 中的分量组成的. 显然 $(x_1^* - x_2^*)_{J_k} \neq 0$, 上式推出 $(aQ)_{J_k \times J_k}$ 是奇异的,此与条件 C 相矛盾. 所以引理 2 结论成立. 至于条件 C_1, C_2, C_3, C_4 成立的情形,我们只需验证它们都蕴含条件 C 成立. 事实上,因为任一主子矩阵的模都不会超过矩阵本身的模,若矩阵 $I-Q$ 的模小于 1,则 Q 是非奇异的. 所以 C_1 蕴含条件 C. 对于 C_2 ,可以利用矩阵的盖尔园性质验证,此时 Q 的任意主子矩阵都没有零特征值,当然是非奇异的. 另外,直接依据 M 矩阵与正定矩阵的性质,可验证 C_3, C_4 均蕴含条件 C 成立. (证毕)

引理 3 模型(1)存在严格的能量函数.

证明 记 $y(t) = x(t) - aQx(t) + aq$, 并令 $E(x) = 1/2x^T Qx - x^T q$. 则对任意 $x_0 \in \Omega$, 由引理 1 知, $x(t, x_0) \in \Omega$, 从而对轨线上的任一点 x 有 $x = P_\Omega(x)$. 对 $E(x)$ 沿轨线 $x(t, x_0)$ 的求导可得:

$$\begin{aligned} \frac{dE(x)}{dt} &= \langle \nabla E(x), \frac{dx}{dt} \rangle \\ &= \langle a(Qx - q), a^{-1}\tau^{-1}(-x + P_\Omega(x - aQx + aq)) \rangle \\ &= \langle a^{-1}\tau^{-1}(P_\Omega(x - aQx + aq) - P_\Omega(x)), -(x - aQx + aq) + x \rangle \\ &= - \sum_{i=1}^N \frac{1}{\alpha_i \tau_i} (g_i(y_i) - g_i(x))(y_i - x_i) \\ &\leq - \min_i \frac{1}{\alpha_i \tau_i} \| P_\Omega(x - aQx + aq) - P_\Omega(x) \|^2 \\ &= - \min_i \frac{1}{\alpha_i \tau_i} \| P_\Omega(x - aQx + aq) - x \|^2 \\ &\leq 0 \end{aligned}$$

上式即说明: $E(x)$ 是系统(1)的一个严格能量函数. (证毕)

由上述引理 1—引理 3 和定理 1, 我们立即可得下述有关模型(1)的基本收敛性定理:

定理 2 当 Q 满足条件 C 或 $(C_1 - C_4)$, 模型 1 是全局收敛的.

定理 2 的意义在于: 它根据问题约束集本身的特性, 给出了模型(1)全局收敛的一般性条件. 它不仅包含了已有的有关 Q 正定情形模型(1)的收敛性结论, 而且它也给出了 Q 不定情形下(既非正定, 也非负定)模型(1)的收敛性的新结论. 然而, 值得注意的是, 定理 2 不适用于 Q 为半正定(即 $E(x)$ 为一般凸函数)的情形. 我们将在下一节对此作专门分析.

4 半正定情形下模型(1)的收敛性分析

这一节中我们将证明如下基本定理.

定理 3 假设矩阵 Q 半正定, 则模型

$$\tau \frac{dx(t)}{dt} = -x(t) + P_{\alpha}(x(t)) - \alpha Qx(t) + \alpha q \quad (4.1)$$

对任何 $\tau > 0$ 是全局收敛的.

证明 由引理 3, 系统 (4.1) 有一个严格的能量函数, 因此系统的极限点都是平衡点. 设 x^* 是轨线的任一极限点, 即存在子列 $\{t_i\}$ 使得 $\lim_{t_i \rightarrow \infty} x(t_i, x_0) = x^*$. $x^* = P_{\alpha}(x^* - \alpha Qx^* + \alpha q)$. 记 $y^* = x^* - \alpha Qx^* + \alpha q$, $y(t) = x(t) - \alpha Qx(t) + \alpha q$, 利用微分方程理论则有:

$$\begin{aligned} \tau \frac{d(x(t) - x^*)}{dt} &= -(x(t) - x^*) + (P_{\alpha}(y) - P_{\alpha}(y^*)) \\ x(t) - x^* &= e^{-\tau^{-1}(t-s)}(x(s) - x^*) + \tau^{-1}e^{-\tau^{-1}t} \int_s^t e^{\tau^{-1}\eta} (P_{\alpha}(y(\eta)) - P_{\alpha}(y^*)) d\eta \\ \|x(t) - x^*\| &\leq e^{-\frac{(t-s)}{\tau}} \|x(s) - x^*\| + \frac{e^{-t/\tau}}{\tau} \int_s^t e^{\frac{\eta}{\tau}} \|P_{\alpha}(y(\eta)) - P_{\alpha}(y^*)\| d\eta \end{aligned} \quad (4.2)$$

下面我们证明 $\lim_{t \rightarrow \infty} \|P_{\alpha}(y(t)) - P_{\alpha}(y^*)\| = 0$.

$$\text{考虑 } E(y(t)) = \sum_{i=1}^N \tau \int_{y_i^*}^{y_i} [g_i(s) - g_i(y_i^*)] ds$$

$g_i(y)$ 是一个非降的函数, 显然 $E(y^*) = 0$, $E(y(t)) \geq 0$, 且连续有下界. 计算 $E(y)$ 关于系统 (4.1) 的导数可得:

$$\begin{aligned} \frac{dE(y(t))}{dt} &= \langle \nabla E(y), \frac{dy}{dx} * \frac{dx}{dt} \rangle \\ &= \langle \tau(P_{\alpha}(y) - P_{\alpha}(y^*)), (I - \alpha Q)\tau^{-1}(-x + P_{\alpha}(y)) \rangle \\ &= \langle P_{\alpha}(y) - P_{\alpha}(y^*), \tau(I - \alpha Q)\tau^{-1}(-x + P_{\alpha}(y)) \rangle \\ &= -\langle P_{\alpha}(y) - P_{\alpha}(y^*), (I - \alpha Q)(x - x^* - (P_{\alpha}(y) - P_{\alpha}(y^*))) \rangle \\ &= -\langle P_{\alpha}(y) - P_{\alpha}(y^*), y - y^* \rangle + \langle P_{\alpha}(y) - P_{\alpha}(y^*), (I - \alpha Q)(P_{\alpha}(y) - P_{\alpha}(y^*)) \rangle \\ &\leq -\langle P_{\alpha}(y) - P_{\alpha}(y^*), P_{\alpha}(y) - P_{\alpha}(y^*) \rangle \\ &\quad + \langle P_{\alpha}(y) - P_{\alpha}(y^*), (I - \alpha Q)(P_{\alpha}(y) - P_{\alpha}(y^*)) \rangle \\ &= -\langle P_{\alpha}(y) - P_{\alpha}(y^*), \alpha Q(P_{\alpha}(y) - P_{\alpha}(y^*)) \rangle \\ &\leq 0 \end{aligned}$$

在上面的推导中, 我们应用了 Q 的半正定性和 $\tau(I - \alpha Q) = (I - \alpha Q)\tau$. 上式说明 $E(y)$ 是单调下降的, 从而极限 $\lim_{t \rightarrow \infty} E(y)$ 存在, 进而由于 x^* 是 $x(t, x_0)$ 的一个极限点, 我们得出 $\lim_{t \rightarrow \infty} E(y(t)) = \lim_{t \rightarrow \infty} E(x(t_i, x_0) - \alpha Qx(t_i, x_0) + \alpha q) = E(y^*) = 0$. 另外, 由 $g_i(x)$ 的定义, 不难验证对任何 i , 都有

$$\int_{x_0}^x (g_i(t) - g_i(x_0)) dt \geq \frac{1}{2}(g_i(x) - g_i(x_0))^2$$

所以由 $E(y)$ 的定义有

$$0 \leq \frac{1}{2}\tau \|P_{\alpha}(y) - P_{\alpha}(y^*)\|^2 \leq E(y)$$

于是我们得出

$$\lim_{t \rightarrow \infty} \|P_{\alpha}(y) - P_{\alpha}(y^*)\|^2 = 0$$

即 $\lim_{t \rightarrow \infty} \|P_{\alpha}(y) - P_{\alpha}(y^*)\| = 0$. 于是存在 $T > 0$, 使得当 $t \geq T$ 时, 有

$$\|P_{\alpha}(y) - P_{\alpha}(y^*)\| \leq \epsilon$$

当 $s \geq T$ 时, 由 (4.2) 式就推出:

$$\|x(t, x_0) - x^*\| \leq e^{-\frac{(t-s)}{\tau}} \|x(s, x_0) - x^*\| + \frac{e^{-t/\tau} * \epsilon}{\tau} \int_s^t e^{\frac{\eta}{\tau}} d\eta$$

$$\leq e^{-\frac{(t-s)}{\tau}} \|x(s, x_0) - x^*\| + \epsilon * [1 - e^{-(t-s)/\tau}]$$

$$0 \leq \limsup_{t \rightarrow \infty} \|x(t, x_0) - x^*\| \leq \epsilon$$

因为 ϵ 可以任意小,

$$\lim_{t \rightarrow \infty} \|x(t, x_0) - x^*\| = 0$$

由极限的唯一性, 这即推出模型(4.1)对任何 $x_0 \in \Omega$ 是收敛的. (证毕)

上述定理 3 将[6]中有关模型(1)对一致凸函数的收敛性推广到了对普通凸函数的收敛性.

5 数值模拟

虽然我们在前两节中证明了模型 1 的全局收敛性, 但网络的平衡点并不一定是优化问题的最优解. 为了澄清这一问题, 我们引入 Xu 等人在文献[1]中提出的正则性, 正规性和完备性的概念如下:

定义 2 设 Ω^* 是优化问题的最优解集, Ω^c 是相应的神经网络模型的平衡点的集合, 若 $\Omega^* \subset \Omega^c$ 则称网络是正规的. 若 $\Omega^* \supset \Omega^c$, 称网络是正则的. 若两者相等, 就称模型是完备的.

根据上述定义, 一个正规的网络可保证其求解优化问题的可靠性(因为它的全局最优解必然在网络的平衡态中); 一个正则的网络保证其求解优化问题的有效性(因为网络的每一个平衡态都给出优化问题的一个解); 而一个完备的网络则可保证该网络既有效又可靠地用于求解优化问题.

依据上述概念, [6]证明了模型(1)对于任意凸函数的完备性以及对于任何非凸函数的正规性. 于是, 从定理 2 与 3 推出: 对于 Q 是正定或半正定的情形, 模型(1)必然能有效可靠的用于求解有界二次规划问题的全局最优解. 而对于 Q 不定情形, 尽管模型(1)是全局收敛的, 但不一定每次必然找到问题的全局最优. 下述例 1 和例 2 分别模拟了以上两种情形.

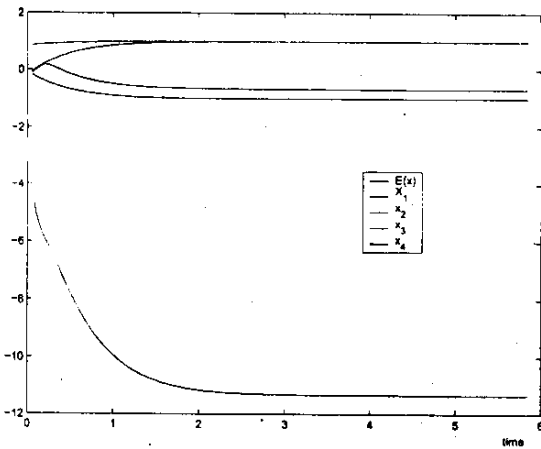


图 1(a)

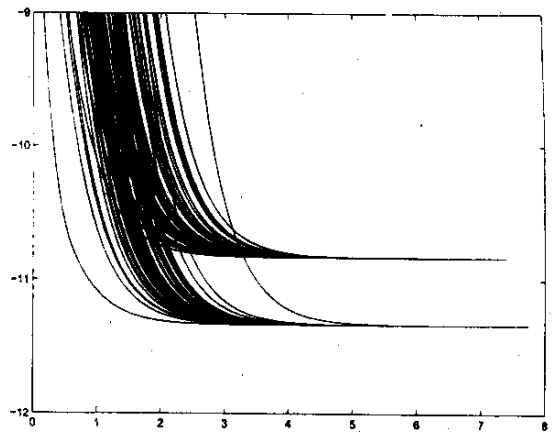


图 1(b)

图 1 不定情形下系统变量和目标函数演化图

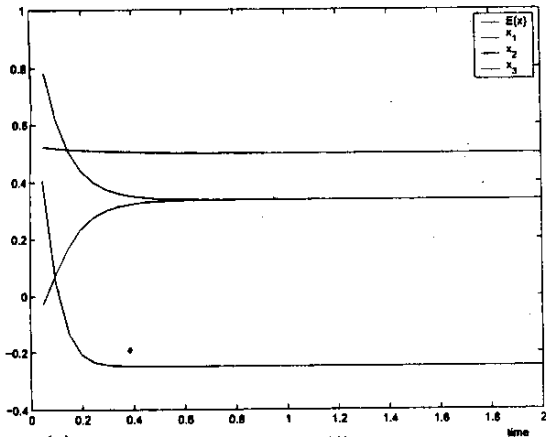
求解问题 $\min E(x) = 1/2x^T Qx - x^T q; x \in \Omega = [-1 \ 1]^n$

例 1 考虑 $Q = \begin{pmatrix} 6 & 8 & 6 & 1 \\ 8 & 2 & 3 & 4 \\ 6 & 3 & 3 & 3 \\ 1 & 3 & 5 & 2 \end{pmatrix}$ 和 $q = (9, 6, 4, 1)^T$ 的情形.

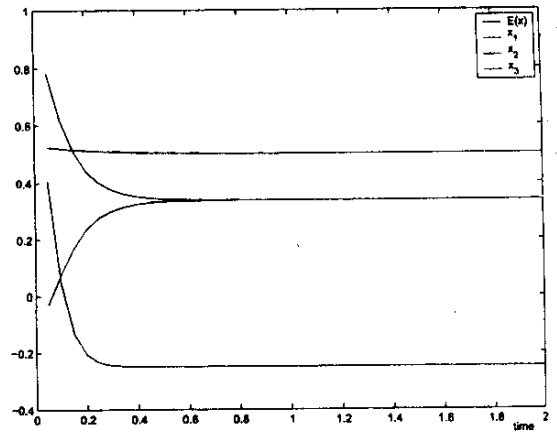
此时, Q 为不定情形. 有两个平衡点 $(-0.6667, 1, 1, -1)^T$ 和 $(1, -1, -0.6667, 1)^T$, 对应的目标函数值分别是 -11.3333 和 -10.8333 . 在图 1(b)中, 通过随机产生 100 个初始点, 可以看到轨线都是收敛的, 这和定理 2 的结论相吻合. 其中, 有 42 次收敛到非最优解的情形. 见图 1(a), 图 1(b).

例2 考虑 $Q = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ 和 $q = (0, 0, 1)^T$ 的情形.

容易验证 Q 是半正定的. 最优解集合 $\{x \mid x_1 = x_2, x_3 = 0.5\}$. 最优目标函数值是 0.25. 随机产生初始点, 观察轨线的演化过程. 试验表明定理 3 的正确性. (图 2) 给出两种运行实例.



(a) The tracks of objective function $E(x)$ and state variables when initial points is $(0.9003, -0.5377, 0.2137)$



(b) The tracks of objective function $E(x)$ and state variables when initial points is $(-0.0280, 0.7628, 0.0242)$

图 2(a)

图 2(b)

图 2 半正定情形下系统变量和目标函数演化图

从以上两个例子可以看出, 对于任何的凸函数 $E(x)$ (即 Q 半正定), 模型(1)确能可靠有效地求出问题的全局最优解. 而对于不定情形, 虽然模型(1)不能百分之百保证求出问题的最优解, 如例 1 中, 有 44% 的可能收敛到非全局最优解, 但仍有相当大的概率求出问题的全局最优解. 这一模拟结果, 不仅与本文的理论分析一致, 也说明了模型(1)确实不失为一种非常好的求解有界二次规划问题的神经网络算法.

6 结论

本文详尽研究了求解有界二次规划问题投影型神经网络的收敛性, 通过深化 Lasella 不变原理, 建立了一个一般性的判别系统全局收敛性的判别准则. 具体应用到模型(1), 我们获得了其全局收敛的一系列充分条件. 特别的, 对于在半正定情形下所获结果推广了近期文献[6, 8]的相应结论. 所有的这些新的结论都表明了模型(1)在求解有界约束二次规划问题的有效性. 模拟实例也表明了这一点.

参考文献:

- [1] Xu Z B, Hu G Q, Kwong C P. Asymmetric hopfield-type networks; theory and applications[J]. Neural Networks, 1996, 9(3): 483-501.
- [2] Xu Z B, Kwong C P. Associative memories[A], ed. by C. T. Leondes, Neural Networks; Techniques and Applications[C]. New York: Academic Press, 1998. 183-258.
- [3] Xia Y S. A new neural network for solving linear and quadratic programming problems[J]. IEEE Transactions on Neural Networks, 1996, 7(6): 1544-1547.
- [4] Zhang X S, Constantinides A G. Lagrange programming neural network[J]. IEEE Transactions on Circuits and Systems, 1992, 39(7): 441-452.

(下转第 128 页)

上面网络割矩阵中的常数项列全部为 1, 共 16 个, 即得到 16 个割集. 割矩阵下边的数字是各弧所对应的弧容量. 网络割矩阵第一行对应的割集的割容量为 $3+5=8$; 第二行对应的割集的割容量为 $5+2=7$; 以此类推, 可得他们的割容量 8, 7, 11, 13, 9, 7, 16, 9, 10, 15, 11, 6, 14, 13, 5, 10.

取 $\min\{8, 7, 11, 13, 9, 7, 16, 9, 10, 15, 11, 6, 14, 13, 5, 10\}=5$

即网络图的最小割容量为 5, 故该网络图的最大流为 5. 而割容量 5 所对应的为第 19 个方程, 割集是由 V_1 流向 \bar{V}_1 的弧为 $(v_s, v_1), (v_2, v_4)$, 故 V_1 的点集由 $\{v_s, v_2\}$ 组成, \bar{V}_1 的点集由 $\{v_1, v_3, v_4, v_t\}$ 组成.

5 结束语

在实际计算时, 网络中点集 V_1 与点集 \bar{V}_1 内部存在不相关联的点, 此割集的容量可以不计算, 因为这类割集的弧包含其他割集的弧 (指从点集 V_1 到点集 \bar{V}_1 上的弧). 因此计算还可简便. 如上例中点集 V_1 分别为 $\{v_s, v_2, v_3, v_4\}, \{v_s, v_1, v_3, v_4\}, \{v_s, v_3, v_4\}, \{v_s, v_2, v_3, v_4\}, \{v_s, v_2, v_3\}, \{v_s, v_1, v_4\}, \{v_s, v_4\}, \{v_s, v_3\}$, 他们分别对应于方程 (7), (8), (11), (13), (14), (17), (18). 在这些割集中, 点集 V_1 与点集 \bar{V}_1 内部没有关联的点, 因此在计算时可以不考虑这些方程.

参考文献:

- [1] 甘应爱, 田丰, 等. 运筹学[M]. 北京: 清华大学出版社, 1999.
- [2] 邹豪思, 王远志. 网络最大流的矩阵算法[J]. 内蒙古大学学报, 2001, 32(4): 466-469.
- [3] 宁宣熙. 网络最大流的图单纯形解法[J]. 南京, 南京航空航天大学学报, 1996, 28(5): 626-630.
- [4] 卢向华, 侯定丕, 魏权龄. 运筹学教程[M]. 北京, 高等教育出版社, 1991.

(上接第 104 页)

- [5] Wu X Y, et al. A high-performance neural network for solving linear and quadratic programming problems[J]. IEEE Transactions on Neural Networks, 1996, 7(3): 643-651.
- [6] Liang X B, Wang J. A recurrent neural network for nonlinear optimization with a continuously differentiable objective function and bound constraints[J]. IEEE Transactions on Neural Networks, 2000, 11(6): 1251-1262.
- [7] Liang X B. A complete proof of global exponential convergence of neural network for quadratic optimization with bound constraints[J]. IEEE Transactions on Neural Networks, 2001, 12(3): 636-639.
- [8] Bouzerdoum A, Pattison T R. Neural network for quadratic optimization with bound constraints[J]. IEEE Transactions on Neural Networks, 1993, 4(2): 293-303.
- [9] Xia Y S, Wang J. On the stability of globally projected dynamical systems[J]. Journal of Optimization Theory and Applications, 2000, 106: 129-150.
- [10] Xia Y S, Wang J. A general methodology for designing globally convergent optimization neural networks[J]. IEEE Transactions on Neural Network, 1998, 9(6): 1331-1343.
- [11] Zhang X S. Neural Network in Optimization[M]. Kluwer Academic Publisher, 2000.

求解有界约束二次规划问题神经网络模型的收敛性分析

作者: 李有梅, 彭济根, 徐宗本
作者单位: 李有梅(西安交通大学理学院信息与系统科学研究所, 陕西, 西安, 710049; 山西大学计算机科学系, 山西, 太原, 030006), 彭济根, 徐宗本(西安交通大学理学院信息与系统科学研究所, 陕西, 西安, 710049)
刊名: 系统工程理论与实践 ISTIC EI PKU
英文刊名: SYSTEMS ENGINEERING—THEORY & PRACTICE
年, 卷(期): 2003, 23(9)
被引用次数: 0次

参考文献(11条)

1. Xu Z B, Hu G Q, Kwong C P Asymmetric hopfield-type networks: theory and applications 1996(03)
2. Xu Z B, Kwong C P Associative memories 1998
3. Xia Y S A new neural network for solving linear and quadratic programming problems 1996(06)
4. Zhang X S, Constantinides A G Lagrange programming neural network 1992(07)
5. Wu X Y A high-performance neural network for solving linear and quadratic programming problems 1996(03)
6. Liang X B, Wang J A recurrent neural network for nonlinear optimization with a continuously differentiable objective function and bound constraints 2000(06)
7. Liang X B A complete proof of global exponential convergence of neural network for quadratic optimization with bound constraints 2001(03)
8. Bouzerdoum A, Pattison T R Neural network for quadratic optimization with bound constraints 1993(02)
9. Xia Y S, Wang J On the stability of globally projected dynamical systems 2000
10. Xia Y S, Wang J A general methodology for designing globally convergent optimization neural networks 1998(06)
11. Zhang X S Neural Network in Optimization 2000

相似文献(10条)

1. 学位论文 陈增强 广义控制系统设计的若干问题研究 1997

该文首先对国内外近年来在预测控制理论及应用方面的研究进展进行了深入的分析与总结,在此基础上选择了预测控制系统设计中的一些重要问题进行了细致的研究。

2. 期刊论文 王海军, 白玫, 贾兆立, 覃丽萍, WANG Hai-jun, BAI Mei, JIA Zhao-li, QIN Li-ping 基于二次优化BP神经网络的期货价格预测—数学的实践与认识 2008, 38(5)

针对BP算法存在的不足,结合神经网络、遗传算法和主成分分析的优点,提出基于二次优化BP神经网络的期货价格预测算法。初次优化采用主成分分析法对网络结构进行优化,第二次优化采用自适应遗传算法对网络参数进行优化,将经过二次优化后建立的BP神经网络模型用于期货价格预测。经仿真检验,用新方法建立的模型对期货价格进行预测,在预测的精度和速度方面都优于单纯BP神经网络模型。

3. 学位论文 王亚赫 基于Hopfield神经网络的盲均衡 2007

无线信道盲均衡是近年来通信信号处理领域的研究热点,现代通信系统中,由于信道的非理想特性而引起的码间干扰是影响通信质量的一个主要因素。传统均衡器克服码间干扰的方法是发送训练序列,盲均衡是一种不借助于训练序列而仅仅利用接收序列的新兴均衡技术。

本论文共分四章。

第一章概述了盲均衡常识:

第二章首先阐述了人工神经网络基础知识,再过渡到神经网络和盲均衡算法相结合,推导出了基于多层前馈神经网络(MLFNN)的盲均衡算法;

第三章深入研究了一些经典二阶统计量算法,并分析了这些传统盲均衡算法的优缺点。

第四章提出了一种基于连续状态离散时间Hopfield神经网络。

盲均衡的新方法:在基于神经网络的能量函数基础上,充分利用接收信号矩阵的正交补空间与发送信号的有限字符集特性,把构造的信号盲检测二次优化形式拟合成具有稳定状态解的连续Hopfield神经网络能量函数形式;结合神经网络自收敛特点与最大程度减少网络平衡点方法,实现对盲信号直接恢复。

计算机仿真表明本文提出的算法表现出了更加良好的性能。

4. 会议论文 宋玉阶, 吴怀宇 求解二次优化问题的神经网络方法 1998

提出了一种求解二次优化问题的神经网络模型, 研究了该神经网络的稳定性和收敛性, 探讨了神经优化的策略, 通过算例证明了该神经网络的可行性。

5. 学位论文 傅莉 反馈神经网络在蜂窝CDMA通信窄带干扰抑制中的应用 1996

该文共分六章, 第一章介绍该课题的研究背景, 包括DS/CDMA蜂窝系统及神经网络的有关知识, 第二章介绍自适应窄带干扰抑制滤波器(Estimator-Subtractor)的数学模型, 比较了线性预测滤波和线性内插滤波的输出误差, 并介绍了两种窄带干扰的模型(一阶AR模型和窄带BPSK模型)。第三章介绍了基于二次优化的Tank-Hopfield神经网络的有关理论, 给出了网络动态方程及网络结构, 该网络适于求解超定线性方程组。作者在第四章中讨论了Tank-Hopfield神经网络在窄带干扰抑制中的应用, 即如何设置网络参数以达到抑制窄带干扰的目的。第五章给出了各种计算机仿真曲线, 作者发现, 滤波网络能在纳秒级内收敛并可大大降低系统误码率。第六章讨论了DS/CDMA蜂窝系统的系统容量, 推导了有关公式并作了计算机仿真工作, 显然, 当信干比(SIR)较低时, 窄带干扰抑制网络能在很大程度上提高系统容量, 最后, 作者对该文作了小结。

6. 期刊论文 宋玉阶, 吴怀宇, Song Yujie, Wu Huaiyu 求解线性约束二次优化问题的神经计算模型 - 计算机工程与科学 2000, 22(2)

本文提出了一种求解线性约束二次优化问题的神经模型, 研究了该神经网络的稳定性和收敛性, 给出了电路框图, 并通过算例证明了该神经网络的可行性。

7. 学位论文 郭永玲 自来水厂出水浊度的智能控制策略研究 2007

本文针对自来水厂的混凝投药过程的复杂性, 分别对先进PID控制、智能控制、二次优化控制技术进行了研究和改进, 探讨各种先进控制算法在混凝投药过程的应用, 具体研究内容如下:

- 1) 对目前控制领域的先进PID控制算法进行了研究, 深入分析常规PID控制器在大时滞、时变等复杂过程控制应用中的局限性, 追究其结构和算法的不足, 从而通过引入先进的算法改善其控制品质。这里着重研究非线性PID控制器和PID神经网络控制器的设计及其控制算法, 并对原有的控制算法作了一些改进, 通过大量的仿真研究, 总结两种先进PID控制算法的特点, 突出改进算法的优越性。
- 2) 对相关不需要精确建模的智能控制算法进行研究。由于影响混凝投药过程控制的外界因素很多, 且不确定, 往往难于对其精确建模。因此, 对相关的智能控制算法进行研究是很必要的。这里主要研究和改进神经网络和预测控制结合的智能控制算法。以神经网络作为预测模型, 进行多步预测, 并充分利用PID神经网络控制器超调小的优点, 研究基于神经网络模型的多步预测控制。
- 3) 对于二次优化控制理论进行了较深入探讨和研究, 这种算法从根本上解决了时滞系统中无穷因子的困扰, 实现了用有限维但不限于三维, 对无穷维对象的控制。由于时间所限, 这里仅对混凝投药控制系统进行初步的设计和仿真研究, 并发现一些应改进的问题。

8. 学位论文 甘艳珍 城市供水处理混凝投药过程的建模与控制研究 2008

混凝剂投加是城市供水处理过程中净水处理的重要环节, 准确地投加混凝剂可以有效地减轻过滤、消毒设备的负担, 在保证满足出厂水浊度要求的前提下尽量减少混凝剂的投加量, 具有良好的经济效益和社会效益。

城市供水处理的混凝投药过程反映到控制理论上是一个大惯性大时滞、受外界因素干扰不确定和难于精确建模的复杂过程。这里针对混凝投药过程的复杂性, 提出了基于神经网络的系统辨识, 得到该水厂混凝投药过程的数学模型, 在此基础上, 分别对基于状态反馈的ITAE控制和神经网络控制进行了研究和改进, 并且通过仿真研究, 验证和比较了各种控制方法的有效性。具体研究内容如下:

- (1) 首先, 通过已知模型的神经网络辨识, 得到了基于神经网络的阶跃响应, 并与对象的实际阶跃响应相对照。仿真结果表明, 可以从神经网络模型中得到被控过程的阶跃响应, 从而得到相应的过程参数。然后, 针对从该水厂获取的工业实测数据, 进行去野值, 去高、低频噪音, 数据平滑处理和归一化等预处理, 得到了可以用于过程辨识的有效数据。再通过分析神经网络的各个输入, 以及根据经验得到各个变量相对应的近似延迟值, 构造了一个比较合理的神经网络模型。通过用经过数据处理的工业数据训练, 得到一个效果较好的神经网络模型。对该神经网络进行阶跃响应实验, 得到基于神经网络的阶跃模型, 从而分析得到水厂混凝投药过程控制通道的模型。这个模型可以为工业上的实际应用提供一定的参考。
- (2) 针对得到的水厂混凝投药过程模型, 把过程的时滞和惯性看做一个整体, 提出了一种基于状态反馈的阶跃响应的ITAE控制研究方法。首先, 采用多项式拟合纯滞后环节, 进而构造基于观测器的内环状态反馈矩阵, 状态反馈矩阵和观测器的极点都配置成具有ITAE最佳传递函数。合理地配置观测器频率和状态频率 ω_0 从而很好地改善了对象动态特性。外环采用基于ITAE最佳设定公式的PID控制器来控制。仿真结果表明: 此种控制方法结合了PID控制和二次优化控制的优点, 对于单容大时滞对象, 具有良好的快速性, 平稳性和鲁棒性。
- (3) 分析了基于BP神经网络的PID控制原理, 系统结构, 以及学习算法。在前人的研究基础上, 提出了基于RBF辨识网络的BP神经网络的PID控制算法。主要从两个方面来改进。1) 通过采用RBF辨识网络得到对象的灵敏度信息 $\delta y(k) / \delta u(k)$; 2) 增加一个参数 α 调整积分作用。大量的仿真研究表明: 此控制方案能有效地抑制超调量, 具有良好的快速性, 平稳性和鲁棒性。

9. 会议论文 宋玉阶, 吴怀宇 一般等式绝症砂二次优化问题的神经计算方法 1998

提出一种求解一般等式约束二次优化问题的神经网络模型, 研究了该神经网络的稳定性和收敛性, 给出了电路实现方法, 通过算例证明了该神经网络的可行性。

10. 学位论文 施昌磊 基于Simulink软件无线电的仿真与研究 2007

本文系统地研究了基于MATLAB中Simulink的通信基带系统建模仿真。利用Simulink对通信系统中几种基本信道模型进行了构造研究, 同时还构造了Hopfield神经网络模型, 并以此解决盲信号检测问题, 仿真结果表明M文件和Simulink模型结果一致。由于Simulink是具有内在的模块化, 系统化, 以及时空流, 定点数的仿真工具, 不仅可以真实地模拟一个通信系统, 使通信系统里面的核心算法在整个系统中进行仿真验证, 而且可以和MATLAB的工具包RTW(Real-Time-Workshop)结合, 产生可以直接在数字信号处理硬件上运行的代码, 所以这些研究工作对于软件无线电技术开发应用具有重要价值。

本文主要工作归纳如下:

1. 利用Simulink对通信系统中最基本信道模型—加性高斯白噪声信道的特性进行仿真。首先介绍了该信道的基本特性, 然后构建了它的系统仿真模型, 介绍了一种基于批处理的仿真系统。
 2. 研究了基带通信系统中Rayleigh衰落信道的特性, 在对Jakes模型进行分析的基础上, 利用Simulink的基础模块搭建了与数学模型对应的完整Rayleigh衰落信道模型。
 3. 学习了解了Hopfield神经网络的相关知识, 由于神经网络的能量函数的特性, 可以利用发送信号的有限字符集特性和接收信号矩阵的正交补空间特性, 把信号盲检测二次优化形式转换为具有稳定状态解的连续Hopfield神经网络能量函数形式, 实现对盲信号的直接恢复, 并用Simulink仿真实现, 验证了其良好的性能。
- 由于Simulink硬件化的特性, 使得本论文的工作有着很大的实用性以及一定的创新性。

本文链接: http://d.g.wanfangdata.com.cn/Periodical_xtgcllys200309017.aspx

授权使用: 西安交通大学(wfxajd), 授权号: bdc16964-2a36-43c0-8755-9db20153e1a1

下载时间: 2010年7月13日