第 49 卷 第 5 期 2010 年 9 月
中山大学学报（自然科学版）
ACTA SCIENTIARUM NATURALIUM UNIVERSITATIS SUNYATSENI
Vol. 49 No. 5 Sep. 2010

一类新的记忆梯度法及其收敛性*

汤京永[1,2]，董 丽[1]
（1. 信阳师范学院数学与信息科学学院，河南 信阳 464000；
2. 上海交通大学数学系，上海 200240）

摘 要：提出一类新的求解无约束优化问题的记忆梯度法。算法在每步迭代中利用当前和前面迭代点的信息产生下降方向，采用精确线性搜索或 Wolfe 非精确线性搜索产生步长，在较弱条件下证明了算法具有全局收敛性和线性收敛速率。数值试验表明算法是有效的。

关键词：无约束优化；记忆梯度法；全局收敛性；线性收敛速率
中图分类号：O221.2 **文献标志码**：A **文章编号**：0529－6579（2010）05－0025－05

A New Memory Gradient Method and Its Convergence

TANG Jingyong[1,2], *DONG Li*[1]
（1. College of Mathematics and Information Science, Xinyang Normal University, Xinyang 464000, China;
2. Department of Mathematics, Shanghai Jiaotong University, Shanghai 200240, China）

Abstract: A new memory gradient method for unconstrained optimization problems is presented. This method makes use of the current and previous iterative information to generate a decent direction and uses exact linear search or Wolfe inexact linear search to define the step-size at each iteration. The global convergence and linear convergence rate are proved under some mild conditions. Numerical experiments show that the new method is efficient in practical computation.

Key words: unconstrained optimization; memory gradient method; global convergence; linear convergence rate

考虑无约束优化问题

$$\min f(x),\ x \in \mathbb{R}^n \tag{1}$$

其中 $\mathbb{R}^n$ 是 n 维欧氏空间，$f(x):\mathbb{R}^n \to \mathbb{R}$ 为连续可微函数。设 $f(x)$ 的梯度函数为 $g(x)$。若 x_k 为当前迭代点，则简记 $\nabla f(x_k)$ 为 g_k，$f(x_k)$ 为 f_k，$f(x^*)$ 为 f^*。

求解问题（1）的算法主要是迭代法，其一般形式为

$$x_{k+1} = x_k + \alpha_k d_k$$

其中 d_k 为下降方向，α_k 为步长参数。对 α_k 和 d_k 的不同选择就构成了不同的迭代算法[1]。一般情况下，共轭梯度法是求解大型无约束优化问题的有效方法之一，它在每步迭代中不需计算和存储矩阵，算法简单，其搜索方向的基本结构为

$$d_k = \begin{cases} -g_k, & k = 1 \\ -g_k + \beta_k d_{k-1}, & k \geq 2 \end{cases}$$

其中 β_k 是一个参数，当 β_k 取不同公式时就得到了不同的共轭梯度法[2-4]。

记忆梯度法类似于共轭梯度法，在每步迭代中不用计算和存储矩阵，算法简单，因而适于求解大规模优化问题，且与共轭梯度法相比，记忆梯度法增加了参数选择的自由度，更有利于构造稳定的快速收敛算[5-12]。文［6］提出一种记忆梯度法，在 Wolfe 步长规则下证明了算法的全局收敛性。文

* 收稿日期：2009－09－11
基金项目：国家自然科学基金资助项目（10571109）；山东省自然科学基金资助项目（Y2008A01）；信阳师范学院青年基金资助项目（200947，200946）
作者简介：汤京永（1979 年生），男，博士，讲师；E-mail：tangjing-yong@ tom. com

[7] 提出一种新的记忆梯度法，在强 Wolfe 步长规则下具有全局收敛性，并且有较好的数值试验结果。

本文提出一类新的求解无约束优化问题的记忆梯度法，算法在每步迭代中利用当前和前面迭代点的信息产生下降方向，采用精确线性搜索或 Wolfe 非精确线性搜索产生步长。新算法有以下优点：①结构简单，不需要计算和存储矩阵，适于求解大规模优化问题；②算法具有全局收敛性和线性收敛速率；③算法具有较好的数值试验结果。

1 记忆梯度法及其性质

算法：取 $0<\rho<1, 0<\sigma_1<\sigma_2<1, x_1\in\mathbb{R}^n$。令 $k:=1$；

步 1：若 $\|g_k\|=0$，则停止迭代；否则，转步 2；

步 2：$x_{k+1}=x_k+\alpha_k d_k$，其中

$$d_k=\begin{cases}-g_k, & k=1\\ -[(1-\beta_k)g_k+\beta_k\lambda_{k-1}], & k\geqslant 2\end{cases}\tag{2}$$

这里

$$\beta_k=\frac{\rho\|g_k\|}{\|g_k\|+\|\lambda_{k-1}\|},\qquad \lambda_{k-1}=g_{k-1}-d_{k-1}\tag{3}$$

而 α_k 由精确线性搜索规则或 Wolfe 步长准则确定。

精确线性搜索规则：求 α_k 使其满足

$$f(x_k+\alpha_k d_k)=\min\{f(x_k+\alpha d_k)\mid\alpha>0\}\tag{4}$$

Wolfe 步长规则：计算 α_k 使其满足

$$f(x_k+\alpha d_k)\leqslant f(x_k)+\sigma_1\alpha g_k^{\mathrm{T}}d_k\tag{5}$$

$$g(x_k+\alpha d_k)^{\mathrm{T}}d_k\geqslant\sigma_2 g_k^{\mathrm{T}}d_k\tag{6}$$

步 3：令 $k:=k+1$，转步 1。

注：在上述算法中，若步长 α_k 由精确线性搜索规则确定，则记该算法为算法（A）；若步长 α_k 由 Wolfe 步长规则确定，则记其为算法（B）。

引理 1 对任意的 $k\geqslant 1$，有 $-g_k^{\mathrm{T}}d_k\geqslant(1-\rho)\|g_k\|^2$。

证明 当 $k=1$ 时，由（2）知 $-g_1^{\mathrm{T}}d_1=\|g_1\|^2\geqslant(1-\rho)\|g_1\|^2$。当 $k\geqslant 2$ 时，由（2）和（3）知

$$\begin{aligned}-g_k^{\mathrm{T}}d_k&=(1-\beta_k)\|g_k\|^2-\beta_k g_k^{\mathrm{T}}\lambda_{k-1}=\\&\|g_k\|^2-\beta_k(\|g_k\|^2+g_k^{\mathrm{T}}\lambda_{k-1})\geqslant\\&\|g_k\|^2-\beta_k(\|g_k\|^2+\|g_k\|\,\|\lambda_{k-1}\|)=\\&\|g_k\|^2-\frac{\rho\|g_k\|}{\|g_k\|+\|\lambda_{k-1}\|}\cdot\\&(\|g_k\|^2+\|g_k\|\cdot\|\lambda_{k-1}\|)=(1-\rho)\|g_k\|^2\end{aligned}$$

引理 2 对任意的 $k\geqslant 1$，有 $\|d_k\|\leqslant(1+\rho)\|g_k\|$。

证明 当 $k=1$ 时，则 $\|d_1\|=\|-g_1\|\leqslant(1+\rho)\|g_1\|$。当 $k\geqslant 2$ 时，因为 $0\leqslant\beta_k<1$，故由（2）（3）知

$$\begin{aligned}\|d_k\|^2&=(1-\beta_k)^2\|g_k\|^2-\\&2(1-\beta_k)\beta_k g_k^{\mathrm{T}}\lambda_{k-1}+\beta_k^2\|\lambda_{k-1}\|^2\leqslant\\&\|g_k\|^2+2\beta_k|g_k^{\mathrm{T}}\lambda_{k-1}|+\beta_k^2\|\lambda_{k-1}\|^2\leqslant\\&\|g_k\|^2+2\beta_k\|g_k\|\,\|\lambda_{k-1}\|+\beta_k^2\|\lambda_{k-1}\|^2=\\&\|g_k\|^2+2\rho\|g_k\|^2\frac{\|\lambda_{k-1}\|}{\|g_k\|+\|\lambda_{k-1}\|}+\\&\rho^2\|g_k\|^2\frac{\|\lambda_{k-1}\|^2}{(\|g_k\|+\|\lambda_{k-1}\|)^2}\leqslant\\&\|g_k\|^2+2\rho\|g_k\|^2+\rho^2\|g_k\|^2=\\&(1+\rho)^2\|g_k\|^2\end{aligned}$$

从而知 $\|d_k\|\leqslant(1+\rho)\|g_k\|$。证毕。

2 算法（A）的全局收敛性

为了下面证明的需要，本文作如下假设：

(H_1)：目标函数 $f(x)$ 在水平集 $L_0=\{x\in\mathbb{R}^n\mid f(x)\leqslant f(x_1)\}$ 上有下界。

(H_2)：梯度函数 $g(x)=\nabla f(x)$ 在包含 L_0 的开凸集 B 上 Lipschitz 连续，即存在常数 $L>0$，使 $\|g(x)-g(y)\|\leqslant L\|x-y\|, \forall x,y\in B$

引理 3[5] 若 (H_1) 和 (H_2) 成立，则对任意 $\alpha>0$ 及 $x_k+\alpha d_k\in B$ 有

$$f(x_k+\alpha d_k)-f(x_k)\leqslant\alpha g_k^{\mathrm{T}}d_k+\frac{1}{2}\alpha^2L\|d_k\|^2$$

定理 1 若 (H_1) 和 (H_2) 成立，$\{x_k\}$ 是由算法（A）产生的迭代点列，则 $\lim\limits_{k\to\infty}\|g_k\|=0$。

证明 令 $\hat{\alpha}_k=\dfrac{-g_k^{\mathrm{T}}d_k}{L\|d_k\|^2}$，则由引理 3 及精确线性搜索规则知

$$\begin{aligned}f_k-f_{k+1}&\geqslant f(x_k)-f(x_k+\hat{\alpha}d_k)\geqslant\\&-\hat{\alpha}g_k^{\mathrm{T}}d_k-\frac{1}{2}\hat{\alpha}^2L\|d_k\|^2=\frac{1}{2L}\left(\frac{g_k^{\mathrm{T}}d_k}{\|d_k\|}\right)^2\end{aligned}\tag{7}$$

从而由引理 1，引理 2 及（7）可知

$$f_k-f_{k+1}\geqslant\frac{1}{2L}\frac{(1-\rho)^2\|g_k\|^4}{\|d_k\|^2}\geqslant\frac{(1-\rho)^2\|g_k\|^2}{2L(1+\rho)^2}\tag{8}$$

因为 $\{f_k\}$ 单调不增且有下界，故知 f_k 有极限，从而由（8）知 $\lim\limits_{k\to\infty}\|g_k\|=0$。

3 算法（B）的全局收敛性

首先，我们将假设 (H_2) 减弱为：

(H_3)：梯度函数 $g(x)$ 在包含 L_0 的开凸集 B 上一致连续。

定理 2　若 (H_1) 和 (H_3) 成立，则算法 (B) 或有限步终止于问题 (1) 的稳定点；或产生无穷点列 $\{x_k\}$，其任意聚点都是问题 (1) 的稳定点。

证明　若 $g(x_k)=0$，则 x_k 为稳定点。假设算法 (B) 产生无穷点列 $\{x_k\}$，x^* 为其任意聚点，则存在子列 $\{x_k:k\in K\}$，$K\subset\{1,2,\cdots\}$，使 $\lim\limits_{k\in K,k\to\infty}x_k=x^*$。下面分两种情况讨论。

(I) $\inf\limits_{k\in K}\alpha_k>0$。

由 (5) 和引理 1 可知

$$f_k-f_{k+1}\geqslant-\sigma_1\alpha_k g_k^{\mathrm{T}}d_k\geqslant\sigma_1(1-\rho)\alpha_k\|g_k\|^2 \tag{9}$$

因为 $\{f_k\}$ 单调不增且有下界，故知 $\{f_k\}$ 有极限，从而由 (9) 知 $\lim\limits_{k\to\infty}\alpha_k\|g_k\|^2=0$，进而可知

$$\lim_{k\in K,k\to\infty}\alpha_k\|g_k\|^2=0 \tag{10}$$

又因为 $\inf\limits_{k\in K}\alpha_k>0$，故由 (10) 可知 $\lim\limits_{k\in K,k\to\infty}\|g_k\|^2=0$，即 $\lim\limits_{k\in K,k\to\infty}\|g_k\|=0$。因为 $\lim\limits_{k\in K,k\to\infty}x_k=x^*$，且 $g(x)$ 连续，所以 $g(x^*)=0$，故知 x^* 是问题 (1) 的稳定点。

(II) $\inf\limits_{k\in K}\alpha_k=0$。

由 $\inf\limits_{k\in K}\alpha_k=0$ 知存在子列 $N\subset K$，使 $\lim\limits_{k\in \mathrm{N},k\to\infty}\alpha_k=0$。因为 $\{x_k:k\in K\}$ 有界，$g(x)$ 连续，故 $\{\|g_k\|:k\in K\}$ 有界，从而由引理 2 可知 $\{\|d_k\|:k\in K\}$ 有界。又因为 $\lim\limits_{k\in \mathrm{N},k\to\infty}\alpha_k=0$，故

$$\lim_{k\in \mathrm{N},k\to\infty}\alpha_k\|d_k\|=0 \tag{11}$$

由 (6) 和引理 1 知

$$\|g(x_{k+1})-g(x_k)\|=\frac{\|g_{k+1}-g_k\|\,\|d_k\|}{\|d_k\|}\geqslant\frac{(g_{k+1}-g_k)^{\mathrm{T}}d_k}{\|d_k\|}\geqslant\frac{(\sigma_2-1)g_k^{\mathrm{T}}d_k}{\|d_k\|}\geqslant\frac{(1-\rho)(1-\sigma_2)\|g_k\|^2}{\|d_k\|}$$

从而由引理 2 知

$$\|g(x_k+\alpha_k d_k)-g_k\|\geqslant\frac{(1-\rho)(1-\sigma_2)\|g_k\|}{1+\rho} \tag{12}$$

因此由 (H_3)，(11) 和 (12) 得 $\lim\limits_{k\in \mathrm{N},k\to\infty}\|g_k\|=0$。又因为 $\lim\limits_{k\in \mathrm{N},k\to\infty}x_k=x^*$，且 $g(x)$ 连续，所以有 $g(x^*)=0$，故知 x^* 是问题 (1) 的稳定点。证毕。

4　线性收敛速率

假设 (H_4)：$f(x)$ 是二次连续可微的一致凸函数。

引理 4[7]　若 (H_4) 成立，则 (H_1) (H_2) 和 (H_3) 均成立，$f(x)$ 在 R^n 上有唯一的极小点 x^*，且存在 $0<m\leqslant M$，使

$$\frac{m}{2}\|x-x^*\|^2\leqslant f(x)-f(x^*)\leqslant\frac{M}{2}\|x-x^*\|^2,\ \forall x\in\mathbb{R}^n \tag{13}$$

$$m\|x-x^*\|\leqslant\|g(x)\|\leqslant M\|x-x^*\|,\ \forall x\in\mathbb{R}^n \tag{14}$$

定理 3　若 (H_4) 成立，$\{x_k\}$ 是由算法 (A) 产生的无穷点列，则 $\{x_k\}$ 至少线性收敛于 x^*。

证明　由 (8) 可知

$$f_k-f_{k+1}\geqslant\frac{(1-\rho)^2\|g_k\|^2}{2L(1+\rho)^2}\geqslant\frac{(1-\rho)^2\|g_k\|^2}{2(L+1)(1+\rho)^2} \tag{15}$$

令 $\omega=\dfrac{(1-\rho)^2}{2(L+1)(1+\rho)^2}$，则知 $0<\omega<\dfrac{1}{2}$，从而由 (15) 知

$$f_k-f_{k+1}\geqslant\omega\|g_k\|^2 \tag{16}$$

因为 x^* 为 $f(x)$ 的唯一极小点，故 $f_k-f^*\geqslant0$，从而由 (13) (14) 和 (16) 知

$$f_k-f_{k+1}\geqslant\omega\|g_k\|^2\geqslant\omega m^2\|x_k-x^*\|^2\geqslant\frac{2\omega m^2}{M}(f_k-f^*)\geqslant\frac{2\omega m^2}{M+m^2}(f_k-f^*)$$

令 $\theta=\dfrac{2\omega m^2}{M+m^2}$，显然 $0<\theta<1$，则

$$f_k-f_{k+1}\geqslant\theta(f_k-f^*) \tag{17}$$

由 (17) 可知

$$f_{k+1}-f^*\leqslant(1-\theta)(f_k-f^*)\leqslant\cdots\leqslant(1-\theta)^k(f_1-f^*)$$

从而由 (13) 知

$$\|x_{k+1}-x^*\|^2\leqslant\frac{2}{m}(f_{k+1}-f^*)\leqslant(1-\theta)^k\frac{2(f_1-f^*)}{m}$$

令 $p=\sqrt{\dfrac{2(f_1-f^*)}{m}}$，$q=\sqrt{1-\theta}\,(0<q<1)$，则有 $\|x_{k+1}-x^*\|\leqslant pq^k$，从而知 $\{x_k\}$ 至少线性收敛于 x^*。证毕。

定理 4　若 (H_4) 成立，$\{x_k\}$ 是由算法 (B) 产生的无穷点列，则 $\{x_k\}$ 至少线性收敛于 x^*。

证明　由 (5) 和引理 1 知

$$f_k-f_{k+1}\geqslant-\sigma_1\alpha_k g_k^{\mathrm{T}}d_k\geqslant\sigma_1(1-\rho)\alpha_k\|g_k\|^2 \tag{18}$$

由（12）和（H_2）知

$$2L\alpha_k \| d_k \| \geqslant \| g(x_k + \alpha_k d_k) - g(x_k) \| \geqslant \frac{(1-\rho)(1-\sigma_2)\| g_k \|}{1+\rho}$$

从而由引理2可知

$$\alpha_k \geqslant \frac{(1-\rho)(1-\sigma_2)\| g_k \|}{2L(1+\rho)\| d_k \|} \geqslant \frac{(1-\rho)(1-\sigma_2)}{2(L+1)(1+\rho)^2}$$

由（18）和（19）可得

$$f_k - f_{k+1} \geqslant \frac{\sigma_1(1-\sigma_2)(1-\rho)^2}{2(L+1)(1+\rho)^2} \| g_k \|^2 \quad (20)$$

令 $\omega' = \frac{\sigma_1(1-\sigma_2)(1-\rho)^2}{2(L+1)(1+\rho)^2}$，则易知 $0 < \omega' < \frac{1}{2}$，从而由（20）知

$$f_k - f_{k+1} \geqslant \omega' \| g_k \|^2$$

定理余下的证明类似于定理3，在此省略。

5 数值试验

我们选取文［5］中的两个算例对算法（B）进行数值试验，并与 Wolfe 搜索下的共轭梯度法和最速下降法进行比较。本文用 IT 表示算法的迭代次数，NA 表示本文提出的新算法，FR、HS 和 CD 分别表示 FR、HS 和 CD 共轭梯度法[1-2]，SM 表示最速下降法。参数取值为 $\rho = 0.18$，$\sigma_1 = 0.48$，$\sigma_2 = 0.85$。表格中的数字为迭代中的目标函数值，舍入成小数点后有三位有效数字，一维搜索全部采用 Wolfe 搜索。记 3.320×10^2 为 $3.320(2)$，1.173×10^{-7} 为 $1.173(-7)$，其余同。计算结果如下。

例1 $f(x) = (1-x_1)^2 + (1-x_{10})^2 + \sum_{i=1}^{9}(x_i^2 - x_{i+1})^2$，$x_0 = (2,3,\cdots,2,3)^{\mathrm{T}}$，$x^* = (1,\cdots,1)^{\mathrm{T}}$，$f^* = 0$。

表1 例1的数值结果

Table 1 Numerical results of text 1

IT	NA	FR	HS	CD	SM
0	3.320 (2)	3.320 (2)	3.320 (2)	3.320 (2)	3.320 (2)
40	1.173 (−7)	3.171 (−10)	8.938 (−7)	6.560 (−10)	4.177 (−9)
60	4.040 (−12)	2.596 (−12)	1.710 (−8)	6.099 (−13)	1.740 (−13)
80	2.402 (−18)	5.049 (−14)	3.235 (−10)	6.410 (−15)	2.065 (−17)

例2 扩展 Beale 函数

$$f(x) = \sum_{i=1}^{\frac{n}{2}} \{[1.5 - x_{2i-1}(1-x_{2i})]^2 + [2.25 - x_{2i-1}(1-x_{2i}^2)]^2 + [2.625 - x_{2i-1}(1-x_{2i}^3)]^2\}, n = 2,4,\cdots$$

$$x_0 = (-2,\cdots,-2)^{\mathrm{T}},$$

$$x^* = (3,0.5,\cdots,3,0.5)^{\mathrm{T}}, f^* = 0,$$

取 $n = 40$。

表2 例2的数值结果

Table 2 Numerical results of text 2

IT	NA	FR	HS	CD	SM
0	9.914 (3)	9.914 (3)	9.914 (3)	9.914 (3)	9.914 (3)
20	3.500 (−3)	1.261 (−4)	7.185 (−1)	4.427 (−4)	6.086 (−1)
25	4.166 (−6)	2.889 (−6)	5.649 (−1)	7.322 (−5)	3.592 (−1)
30	2.629 (−7)	2.411 (−5)	4.612 (−1)	1.370 (−5)	2.045 (−1)

例3 在例2中取 $n = 80$。

表3 例3的数值结果

Table 3 Numerical results of text 3

IT	NA	FR	HS	CD	SM
0	1.983 (4)	1.983 (4)	1.983 (4)	1.983 (4)	1.983 (4)
20	7.000 (−3)	2.523 (−4)	1.437 (0)	8.853 (−4)	1.217 (0)
25	8.332 (−6)	5.778 (−5)	1.297 (0)	1.464 (−5)	7.184 (−1)
30	5.258 (−7)	4.822 (−5)	9.223 (−1)	2.740 (−6)	4.089 (−1)

从以上数值试验及比较的结果可以看出，本文提出的新算法具有较好的计算效能，特别是当问题的规模较大时，新算法有更好的表现。

参考文献：

［1］ 袁亚湘，孙文瑜．最优化理论与方法［M］．北京：科学出版社，1997．

［2］ 戴彧虹，袁亚湘．非线性共轭梯度法［M］．上海：上海科学技术出版社，2000．

［3］ 白延琴，张连生．关于共轭梯度法的下降性和收敛性［J］．运筹学学报，2000，4(2)：19-26．

[4] LI Z, WEI J Z, DONG H L. Global convergence of a modified Fletcher-Reeves conjugate gradient method with Armijo-type line search [J]. Numer Math, 2006, 104 (4):561 -572.

[5] SHI Z J. A new super-memory gradient method for unconstrained optimization [J]. Advances in Mathematics, 2006, 3(35):265 -273.

[6] 时贞军. Wolfe 搜索下记忆梯度法的收敛性[J]. 应用数学学报, 2006,1(29):9 -18.

[7] 汤京永,时贞军. 一类全局收敛的记忆梯度法及其线性收敛性[J]. 数学进展, 2007,36(1):67 -75.

[8] 汤京永,董丽. 一类新的曲线搜索下的多步下降算法[J]. 应用数学, 2009, 22 (4): 815 -820.

[9] 汤京永,董丽,张秀军. 一类新的 Wolfe 线性搜索下的记忆梯度法[J]. 山东大学学报, 2009, 44(7):33 -37.

[10] 汤京永,董丽,郭淑利. 一类新的曲线搜索下的记忆梯度法[J]. 信阳师范学院学报,2009, 22(2):179 -182.

[11] YASUSHI, NARUSHIMA. Global convergence of a memory gradient method for unconstrained optimization [J]. Computational Optimization and Applications, 2006, 35(3): 325 -346.

[12] MIELE A, CANTRELL J W. Study on a memory gradient method for the minimization of functions [J]. JOTA, 1969, 3(6): 459 -470.

(上接第 20 页)

参考文献:

[1] DAUBECHIES I. Lectures on wavelets [M]. Philadelphia: Society for Industrial and Applied Mathematics, 1992.

[2] DAUBECHIES I. Orthonormal bases of compactly supported wavelets [J]. Comm Pure and Appl Math, 1988, 41: 909 -996.

[3] 关履泰. 小波方法与应用[M]. 北京:高等教育出版社, 2007.

[4] ANDREAS R. The wavelet transform on Sobolev spaces and its approximation properties [J]. Numer Math, 1991,58: 875 -894.

[5] BOWNIK M. The construction of r-regular wavelets for arbitrary dilation [J]. J Fourier Anal Appl, 2001,7: 489 -506.

[6] HAN B. Symmetry property and construction of wavelets with a general dilation matrix [J]. Linear Algebra Appl, 2002, 35:207 -225.

[7] PETUKHOV A. Construction of symmetric orthogonal bases of wavelets and tight wavelet frames with integer dilation factor [J]. Appl Comput Harmon Anal, 2004, 17: 198 -210.

[8] KAROUI A. A note on the design of nonseparable orthonormal wavelet bases of $L^2(\mathbb{R}^3)$ [J]. Applied Mathematics Letters, 2005, 18(3):293 -298.

[9] LI Y Z. On the holes of a class of bidimensional nonseparable wavelets [J]. Journal of Approximation Theory, 2003,125(2): 151 -168.

[10] HE W J, LAI M J. Construction of trivariate compactly supported biorthogonal box spline wavelets [J]. Journal of Approximation Theory, 2003,120:1 -19.

[11] SKOPINA M. On construction of multivariate wavelets with vanishing moments [J]. Appl Comput Harmon Anal, 2006, 20:375 -390.

[12] 龙瑞麟. 高维小波分析[M]. 北京:世界图书出版公司, 1995.

[13] 黄永东,程正兴. 三元正交小波的构造[J]. 应用数学, 2006, 19 (1):176 -182.