

参赛密码 _____
(由组委会填写)



第十届华为杯全国研究生数学建模竞赛

题 目 功率放大器非线性特性及预失真建模

摘 要：

本文以正交基多项式逼近原理、最小二乘拟合等为基础，建立了无记忆功放和有记忆功放的数学模型及其各自的预失真模型，并用归一化均方误差(NMSE)对所建模型进行了评价，评价结果验证了所建模型的准确性和所用方法的可靠性。

为了建立无记忆功放模型，本文利用多项式函数来表达功率放大器的非线性特征，提出基于样本实际分布的正交基多项式模型，通过最小二乘方法拟合得到模型的具体参数。结果表明，正交基多项式模型具有最好的 NMSE (-157dB)，远优于传统多项式模型和基于均匀分布的正交多项式模型。为避免过拟合，分析得到模型的最优阶次为 7，此时对应的 NMSE 为-68.63dB。

为了建立无记忆功放预失真模型，本文采用间接学习结构对预失真模型进行参数估算。经理论分析得到线性化后最大可能放大倍数为 1.8264 倍。利用最小二乘拟合得到预失真器的正交基多项式模型。为避免过拟合，得到预失真器模型的最优阶次为 5，利用此预失真器模型对功放进行补偿后的输出与具有同样放大倍数的理想功放的输出进行比较，得到其 NMSE 为-37.08dB。

为了建立有记忆功放模型，我们采用在和记忆多项式中所计算得到的正交基进行拟合。利用最小二乘法拟合参数。同样为防止过拟合，经分析得到记忆深度为 5，最大阶次为 4 的多项式模型。此时得到的 NMSE 为-45.26dB

有记忆功放预失真模型也利用间接学习结构得到的。通过建立预失真器输出最大幅度与放大倍数的关系曲线，得知最大放大倍数为 9.5。利用最小二乘拟合得到预失真器的正交基多项式模型。为避免过拟合，得到最优预失真器模型的记忆深度为 5，最大阶次为 5。将该模型补偿之后求得的功放输出与放大倍数为 9.5 的理想功放的输出进行比较，得其 NMSE 为-43.40dB。

对于信号的相邻功率信道比(ACPR)及功率谱密度的拓展研究中，经分析

得知传输信道为 $[0, 20\text{MHz}]$ ，相邻信道为 $[20\text{MHz}, 40\text{MHz}]$ 。通过计算有记忆的预失真器模型补偿后功放的输出信号的功率谱，得到该模型的 ACPR 为 -54.34dB 。未经过预失真补偿信号的 ACPR 为 -39.38dB 。获得了 -15dB 的补偿。

本文的创新点在于：1、对数据进行最小二乘拟合时，采用基于样本实际分布的正交多项式基函数，且正交基的建立使得最小二乘算法可获得高精度并稳健的拟合精度。2、在确定所建多项式模型的最优阶数时，考虑过拟合问题，进行折中的参数选取，使模型具有泛化能力。

关键词：功率放大器 预失真 正交基多项式 最小二乘拟合 间接学习结构

目录

1. 问题重述	- 4 -
2. 模型假设与参量设定	- 5 -
2.1 符号定义与说明	- 5 -
2.2 评价指标选择	- 5 -
3. 模型求解与分析	- 7 -
3.1 无记忆功放问题分析	- 7 -
3.2 无记忆功放模型的建立与求解	- 7 -
3.2.1 传统多项式模型	- 7 -
3.2.2 正交基多项式模型	- 9 -
3.2.3 方案比较	- 13 -
3.3 无记忆功放预失真模型的建立与求解	- 14 -
3.3.1 正交多项式预失真模型建立	- 14 -
3.3.2 线性放大倍数 g 的确定	- 16 -
3.3.3 正交多项式预失真模型求解	- 17 -
3.3.4 小结	- 18 -
3.4 有记忆功放问题分析	- 18 -
3.5 有记忆功放模型的建立与求解	- 19 -
3.5.1 有记忆功放模型建立	- 19 -
3.5.2 有记忆功放模型求解	- 19 -
3.6 有记忆功放预失真模型的建立与求解	- 20 -
3.6.1 正交多项式预失真模型	- 20 -
3.6.2 线性放大倍数 g 的确定	- 23 -
3.6.3 有记忆功放预失真模型求解	- 23 -
3.6.4 小结	- 24 -
3.7 相邻功率信道比及功率谱密度	- 24 -
3.7.1 问题分析	- 24 -
3.7.2 三类信号的功率谱密度求解	- 25 -
3.7.3 结果分析	- 26 -
4 模型改进	- 27 -
参考文献	- 27 -

1. 问题重述

信号的功率放大是电子通信系统的关键功能之一，其实现模块称为功率放大器（PA，Power Amplifier），简称功放。功放的输出信号相对于输入信号可能产生非线性变形，这将带来无益的干扰信号，影响信息的正确传递和接收，此现象称为非线性失真。线性化技术可以在不降低线性度的情况下提高功放的效率，其中数字预失真技术是一种先进的线性化技术。预失真的基本原理是：在功放前设置一个预失真处理模块，这两个模块的合成总效果使整体输入-输出特性线性化，输出功率得到充分利用，其原理框图如图 1 所示。

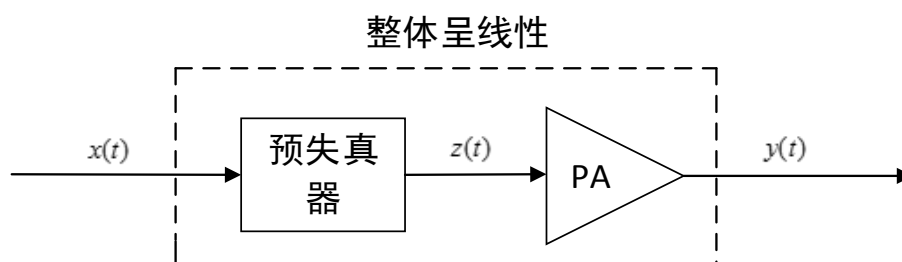


图1 预失真技术的原理框图

图 1 中， $x(t)$ 为输入信号， $y(t)$ 为输出信号， $z(t)$ 为预失真器的输出。

预失真器的非线性特性可描述为 $z(t) = F(x(t))$ ，功放的非线性特性描述为 $y(t) = G(z(t))$ ，预失真处理后整体的线性特性可表示为： $y(t) = G(F(x(t))) = g \cdot x(t)$ ，其中， g 为系统整体的线性放大倍数。

根据某一时刻的输入是否与此前时间段内的输入有关，可将功放分为无记忆功放和有记忆功放两种类型。在保证工程的可实现性并考虑适当的计算复杂度的情况下，针对这两种类型的功放进行以下问题的研究：

问题一：

- A. 建立无记忆功放非线性特性的数学模型，并评价所建模型的准确度。
- B. 设计符合线性化原则以及“输出幅度限制”和“功率最大化”约束的无记忆功放的预失真模型。给出目标误差函数，计算线性化后最大可能的幅度放大倍数，并评价预失真补偿的结果。

问题二：

- A. 建立此功放的非线性特性的数学模型，并评价所建模型的准确度。
- B. 根据线性化原则以及“输出幅度限制”和“功率最大化”约束，以框图的方式建立预失真处理的模型实现示意图，计算预失真模型相关参数，并评价预失真补偿的计算结果。

问题三：

计算功放预失真补偿前后的功率谱密度，并用图形的方式表示三类信号的功率谱密度（输入信号、无预失真补偿的功率放大器输出信号、采用预失真补偿的功率放大器输出信号），并对结果进行分析评价。

2. 模型假设与参量设定

2.1 符号定义与说明

表1 文中主要参量符号定义

符号字母	符号定义
$x()$	预失真器的输入信号
$y()$	功率放大器的输出信号
$z()$	预失真器的输入信号兼功率放大器的输出信号
$\hat{z}()$	预失真训练模块计算的输出值
$e()$	预失真器输出和预失真训练模块之间的估计误差
g	功放的线性放大倍数
ρ	矩阵的条件数
K	多项式的阶数
M	有记忆功放的记忆深度
N	采样样本的个数
Ψ	正交基
J	目标误差函数
O	过拟合程度
NMSE	归一化均方误差
EVM	误差矢量幅度
ACPR	相邻信道功率比

2.2 评价指标选择

对于所求模型的数值计算结果，业界常用 NMSE、EVM 等参数评价其准确度。归一化均方误差 (Normalized Mean Square Error, NMSE)，其表达式定义为：

$$NMSE = 10 \log_{10} \frac{\sum_{n=1}^N |z(n) - \hat{z}(n)|^2}{\sum_{n=1}^N |z(n)|^2} \quad (1)$$

如果用 z 表示实际信号值， \hat{z} 表示通过模型计算的信号值，NMSE 就反映了模型与物理实际模块的接近程度。功放前加载预失真处理后，也可用 NMSE 判断整体模型输出值与理想输出值的近似程度。

误差矢量幅度 (Error Vector Magnitude, EVM) 定义为误差矢量信号平均功率的均方根和参照信号平均功率的均方根的比值，以百分数形式表示。如果用 x 表示理想的信号输出值， e 表示理想输出与整体模型输出信号的误差，可用 EVM 衡量整体模型对信号的幅度失真程度，其表达式如下：

$$EVM = \sqrt{\frac{E[e^2]}{E[X^2]}} \times 100\% \quad (2)$$

分析上述两个指标，当 $z(n)$ 为复数时，有 $|e^2(n)| = |z(n) - \hat{z}(n)|^2$ ，因此当 NMSE 最小时，可知 $|z(n) - \hat{z}(n)|^2$ 最小，从而 $E[e^2]$ 一定最小，则 EVM 也一定最小，这

说明满足 NMSE 指标要求时，必定满足 EVM 要求，因此，我们后续问题的评价指标只计算 NMSE 的值。

3. 模型求解与分析

3.1 无记忆功放问题分析

所谓无记忆功放就是认为放大器的输出信号只取决于当前的输入信号，而与历史输入信号无关。在建立无记忆功放的非线性模型时，通常采用是多项式模型，即泰勒（Taylor）多项式模型。假设功放的输入信号为 $z(t)$ ，那么功放的输出可以表示为：

$$y(t) = \sum_{k=1}^K a_k z^k(t) \quad (3)$$

其中， a_k 为多项式的系数， K 是多项式函数的最大阶数，阶次 K 越大，则模型的准确度越高。

考虑到后面计算中的存在求逆问题，换用一种正交基的多项式结构模型，它能够很好地对功放进行建模，并且减小了传统多项式模型的数值稳定性问题造成的影响。因此，对于问题一，我们的求解思路如下：

(1) 对无记忆功放的建模分别采用两种方案求解，即：传统多项式模型与正交基多项式模型。在正交基多项式模型中，改进了参考文献[3]里的均匀分布情况下的正交基函数，利用数据文件1构建了符合实际分布的正交基函数。利用最小二乘拟合求解出各种方案的模型及其NMSE，并对结果进行了比较分析。最后考虑模型的泛化性能，获得折中时该模型的具体参数。

(2) 对预失真模型的建立采用数据实际分布情况下的正交基多项式形式；分析出“输出幅度限制”和“功率最大化”两个约束条件下可能的线性化的最大幅度放大倍数；采用自适应间接学习结构估算预失真模型的参数，写出其目标误差函数；用最小二乘算法拟合求解出预失真模型及其重构时的NMSE。综合该NMSE及模型的总体泛化性能，取得折中模型参数。最后通过评价经过预失真矫正后的信号与理想线性功放之间的NMSE来对模型进行评价。

3.2 无记忆功放模型的建立与求解

3.2.1 传统多项式模型

分析题目所给的数据文件1，可知无记忆功放的输出幅度存在非线性失真，而输出相位则保持线性，如图2(a)、(b)所示。

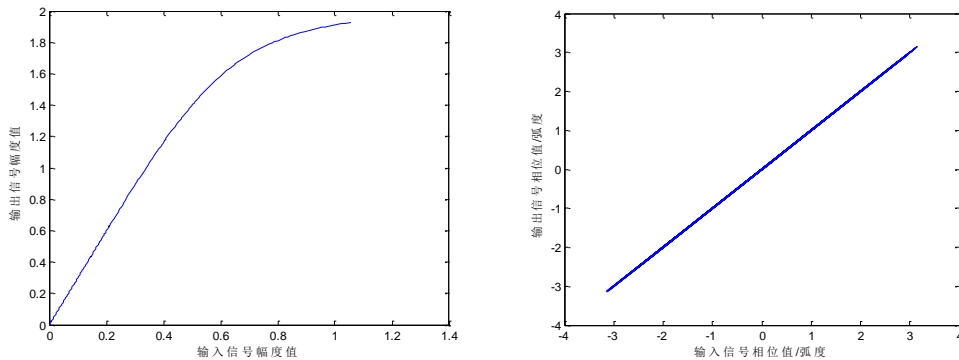


图2 (a) 无记忆功放输入/输出的幅度特性 (b) 无记忆功放输入/输出的相位特性

因此本模型采用的多项式结构如式 (4) 所示：

$$y(t) = \sum_{k=1}^K a_k z(t) |z(t)|^{k-1} \quad (4)$$

将式（4）离散化后得到：

$$y(n) = \sum_{k=1}^K a_k z(n) |z(n)|^{k-1} \quad (5)$$

为增加模型的准确度，这里同时考虑输入信号的奇次项与偶次项，且多项式中不存在常数项。

令 $\Phi_k(x(n)) = x(n)|x(n)|^{k-1}$ ，则功放的输出信号可以写为：

$$y(n) = \sum_{k=1}^K a_k x(n) |x(n)|^{k-1} = \sum_{k=1}^K a_k \Phi_k(x(n)) \quad (6)$$

定义输入和输出向量： $\mathbf{x} = [x(1), \dots, x(n)]^T$, $\mathbf{y} = [y(1), \dots, y(n)]^T$ ，定义参数向量： $\mathbf{a} = [a_1, a_2, \dots, a_K]^T$ ，定义 $N \times 1$ 阶向量

$\Phi_k(x) = [\Phi_k(x(1)), \dots, \Phi_k(x(n))]^T$ ， $N \times K$ 矩阵 $\Phi = [\Phi_1(\mathbf{x}), \dots, \Phi_K(\mathbf{x})]$ 则式（6）写成向量形式：

$$\mathbf{y} = \Phi \mathbf{a} \quad (7)$$

那么，多项式的系数可由最小二乘估计得到：

$$\hat{\mathbf{a}}_{LS} = [\hat{a}_1, \dots, \hat{a}_K]^T = (\Phi^T \Phi)^{-1} \Phi^T \mathbf{y} \quad (8)$$

根据上述理论，我们用 MATLAB2012 编程实现最小二乘算法，拟合得到 $K=1,3,5,7$ 时的无记忆功放的各阶非线性数学模型，并根据公式（1）得到各阶多项式模型的准确度，计算结果如表 2 所示，效果图如图 3 所示。

表2 传统各阶次多项式模型及其准确度

阶次 K	数学模型表达式	NMSE
1	$y(n) = 2.50z(n)$	-16.35
3	$y(n) = 3.05z(n) + 0.01z(n) z(n) - 1.17z(n) z(n) ^2$	-37.81
5	$y(n) = 2.97z(n) + 0.31z(n) z(n) - 0.15z(n) z(n) ^2 - 3.42z(n) z(n) ^3 + 2.21z(n) z(n) ^4$	-53.44
7	$y(n) = 3.01z(n) - 0.28z(n) z(n) + 1.94z(n) z(n) ^2 - 4.6z(n) z(n) ^3 - 3.21z(n) z(n) ^4 + 8.84z(n) z(n) ^5 - 3.80z(n) z(n) ^6$	-68.63

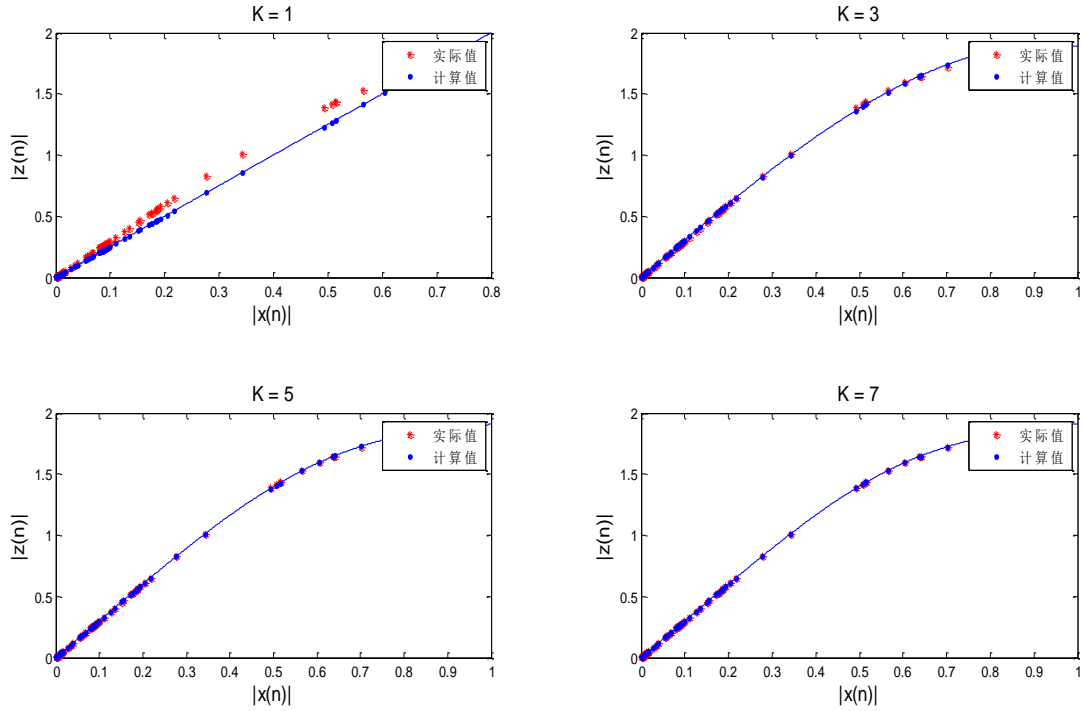


图 3 无记忆功放各阶多项式模型的效果图

从计算结果与效果图中均可看出，随着阶次的增大，多项式模型的准确度越来越好。但由前面的推导可知，我们在用最小二乘法对多项式模型进行参数估计的时候，需要求矩阵 $\Phi^T \Phi$ 的逆矩阵 $(\Phi^T \Phi)^{-1}$ ，在求逆矩阵的时候会遇到数值稳定性问题。定义矩阵的条件数 $\rho = \lambda_{\max} / \lambda_{\min}$ ，其中 λ_{\max} 和 λ_{\min} 分别是矩阵的最大特征值和最小特征值，矩阵的条件数表示矩阵的计算对误差的敏感性，条件数越大，数值稳定性越差，当条件数远大于 1 时，数值误差将非常大。通常，随着多项式阶次的增加，矩阵的条件数也迅速增加，这意味着我们求 $\Phi^T \Phi$ 的逆矩阵 $(\Phi^T \Phi)^{-1}$ 时非常困难，参数估计存在不稳定性。

3.2.2 正交基多项式模型

为缓解数值稳定性问题，我们引入一种正交基函数的多项式模型。把式 (7) 改写成：

$$\mathbf{y} = \boldsymbol{\psi} \boldsymbol{\beta} \quad (9)$$

式中， $\boldsymbol{\psi} = [\psi_1(\mathbf{z}), \psi_2(\mathbf{z}), \dots, \psi_K(\mathbf{z})]$ ，是一组正交基。

用最小二乘估计的表示为：

$$\hat{\boldsymbol{\beta}}_{LS} = (\boldsymbol{\psi}^T \boldsymbol{\psi})^{-1} \boldsymbol{\psi}^T \mathbf{y} \quad (10)$$

则要求解式 (10)，首先要确定正交多项式的基函数 $\boldsymbol{\psi}$ ，此处，我们选用上三角矩阵 \mathbf{U} 来构造正交多项式基函数，矩阵 \mathbf{U} 的第 (k, l) 个元素为 U_{lk} ，所构造的正交多项式基为：

$$\boldsymbol{\psi} = \Phi \mathbf{U} \quad (11)$$

它使 $E[\boldsymbol{\Psi}^T \boldsymbol{\Psi}] = \mathbf{U}^T E[\Phi^T \Phi] \mathbf{U}$ 为对角矩阵。

我们首先按照参考文献[3]，假设输入幅度为均匀分布，令 $r = |x|$ 是 $(0, 1)$ 的均匀分布，得到矩阵 \mathbf{U} 的表达式为：

$$U_{lk} = \begin{cases} (-1)^{l+k} \frac{(k+l)!}{(l-1)!(l+1)!(k-l)!}, & l \leq k \\ 0, & l > k \end{cases} \quad (12)$$

则对于均匀分布 $r=|z|, 2$ 取参考文献[3]其 $K=7$ 阶正交基函数见表 3 所示。

表 3 均匀分布情况的正交基函数

$\psi_1(z) = z$
$\psi_2(z) = 4 z z - 3z$
$\psi_3(z) = 15 z ^2z - 20 z z + 6z$
$\psi_4(z) = 56 z ^3z - 105 z ^2z + 60 z z - 10z$
$\psi_5(z) = 210 z ^4z - 504 z ^3z + 420 z ^2z - 140 z z + 15z$
$\psi_6(z) = 792 z ^5z - 2310 z ^4z + 2520 z ^3z - 1260 z ^2z + 280 z z - 21z$
$\psi_7(z) = 3003 z ^6z - 10296 z ^5z + 13860 z ^4z - 9240 z ^3z + 3150 z ^2z - 504 z z + 28z$

那么，我们同样用最小二乘估计，用MATLAB2012编程求解式（10），得到均匀分布情况的正交基多项式的系数，从而拟合得到 $K=1,3,5,7$ 时的无记忆功放的各阶非线性数学模型，并根据公式（1）得到各阶多项式模型的准确度，计算结果如表4所示，效果图如图4。

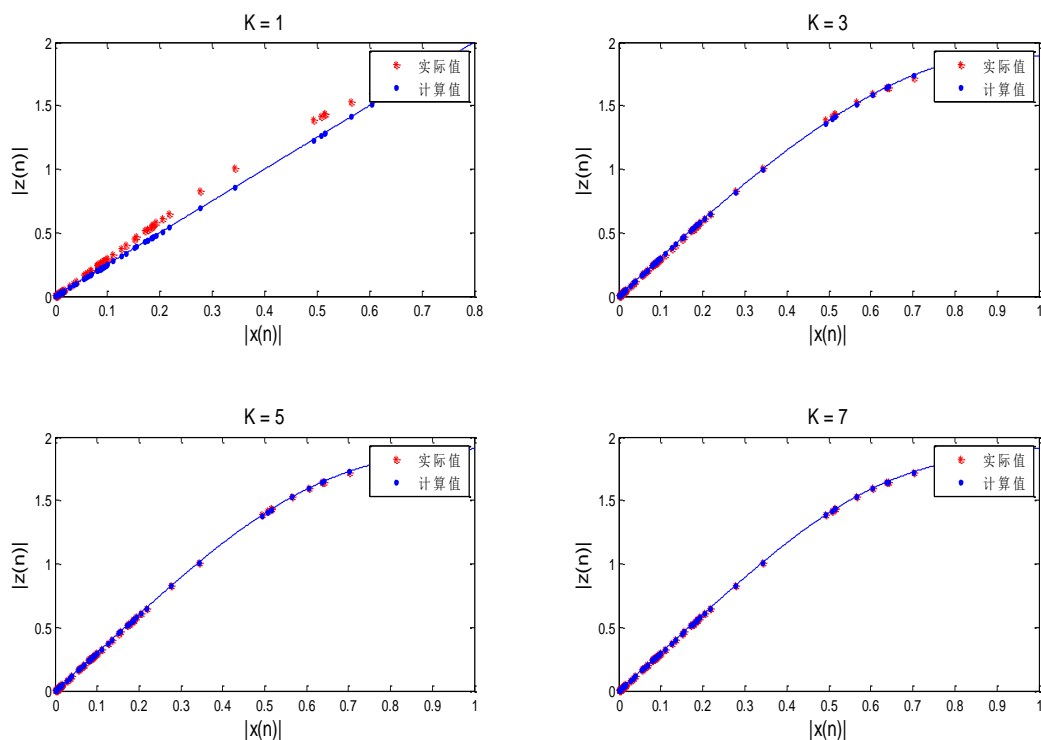


图 4 均匀分布情况的无记忆功放各阶多项式模型拟合效果图

表4 均匀分布情况的各阶正交基多项式模型及其准确度

阶次 K	数学模型表达式	NMSE
1	$y(n) = 2.50z(n)$	-16.34
3	$y(n) = 2.82z(n) - 0.39 z(n) z(n) - 0.16 z(n) ^2 z(n)$	-37.81
5	$y(n) = 3.94z(n) - 0.54 z(n) z(n) - 0.73 z(n) ^2 z(n) + 0.20 z(n) ^3 z(n) + 0.25 z(n) ^4 z(n)$	-53.44
7	$y(n) = 5.46z(n) + 2.26 z(n) z(n) - 3.03 z(n) ^2 z(n) - 0.10 z(n) ^3 z(n) + 2.25 z(n) ^4 z(n) - 0.63 z(n) ^5 z(n) - 0.91 z(n) ^6 z(n)$	-68.63

此计算结果与效果图表明，当 $K \leq 7$ 时，在均匀分布情况下的正交基多项式模型可用来表示无记忆功放的非线性特征，且随着阶次增加，准确度变高。而分析得数据文件1中输入信号幅度并不是均匀分布，如图5所示。

在此基础上对正交基函数的构造进行了改进，利用数据的实际分布构造新的正交基。按照前述构造正交基的方法，即构造矩阵 \mathbf{U} 使 $E[\Psi^T \Psi] = \mathbf{U}^T E[\Phi^T \Phi] \mathbf{U}$ 为对角阵，从而可以得到由式（11）建立的正交基。令：

$$E(\Phi^T \Phi) = \frac{1}{N} \sum_{n=1}^N \Phi^T(z(n)) \Phi(z(n)) \quad (13)$$

根据数据文件1输入信号幅度的实际分布可求解式（13），而根据相似对角化原理，求解 $E[\Phi^T \Phi]$ 的特征向量矩阵作为矩阵 \mathbf{U} ，从而可由式（11）得到实际分布情况的 $K=7$ 阶正交基函数，如表5所示。

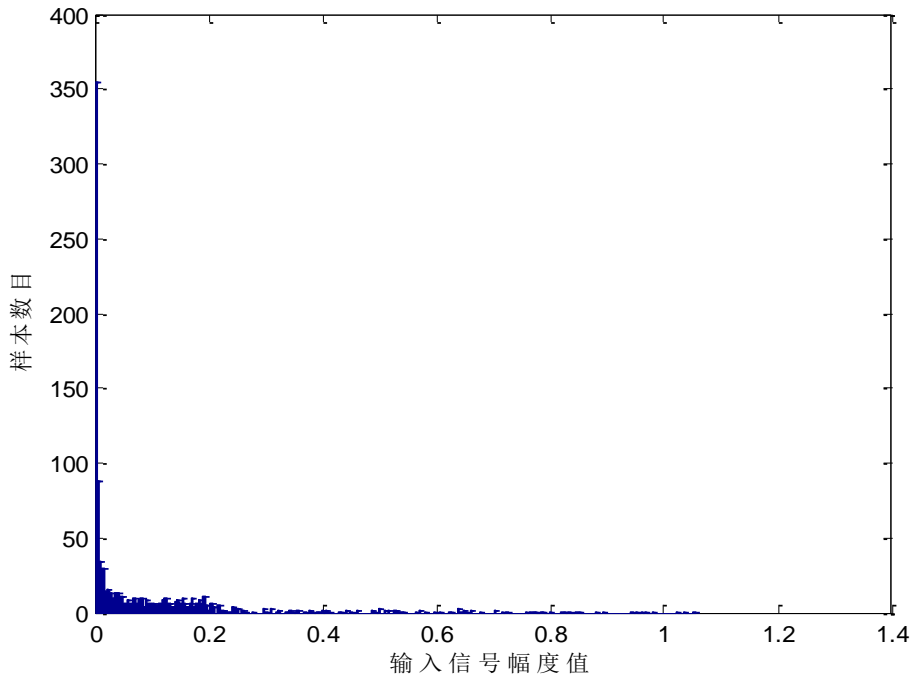


图5 数据文件1输入信号幅度的实际分布

表 5 数据文件 1 分布情况的正交基函数

$\psi_1(z) = 0.15 z ^6 z - 0.53 z ^5 z + 0.69 z ^4 z - 0.44 z ^3 z + 0.14 z ^2 z - 0.02 z z$
$\psi_2(z) = 0.31 z ^6 z - 0.57 z ^5 z - 0.01 z ^4 z + 0.61 z ^3 z - 0.44 z ^2 z + 0.11 z z + 0.01z$
$\psi_3(z) = 0.43 z ^6 z - 0.31 z ^5 z - 0.47 z ^4 z + 0.63 z ^2 z - 0.33 z z + 0.04z$
$\psi_4(z) = -0.51 z ^6 z - 0.06 z ^5 z + 0.31 z ^4 z + 0.48 z ^3 z + 0.24 z ^2 z - 0.59 z z + 0.13z$
$\psi_5(z) = 0.47 z ^6 z + 0.30 z ^5 z + 0.10 z ^4 z - 0.14 z ^3 z - 0.42 z ^2 z - 0.58 z z + 0.38z$
$\psi_6(z) = 0.40 z ^6 z + 0.37 z ^5 z + 0.33 z ^4 z + 0.27 z ^3 z + 0.16 z ^2 z - 0.07 z z - 0.70z$
$\psi_7(z) = 0.26 z ^6 z + 0.27 z ^5 z + 0.29 z ^4 z + 0.32 z ^3 z + 0.37 z ^2 z + 0.44 z z + 0.59z$

那么，我们同样用最小二乘估计：

$$\hat{\delta}_{LS} = (\Psi^T \Psi)^{-1} \Psi^T \mathbf{y} \quad (14)$$

用MATLAB2012编程求解式（14），拟合得到实际分布情况的正交基多项式的系数，从而得到K=1,3,5,7时的无记忆功放的各阶非线性数学模型，并根据公式（1）得到各阶多项式模型的准确度，计算结果如表6所示，效果图见图6。

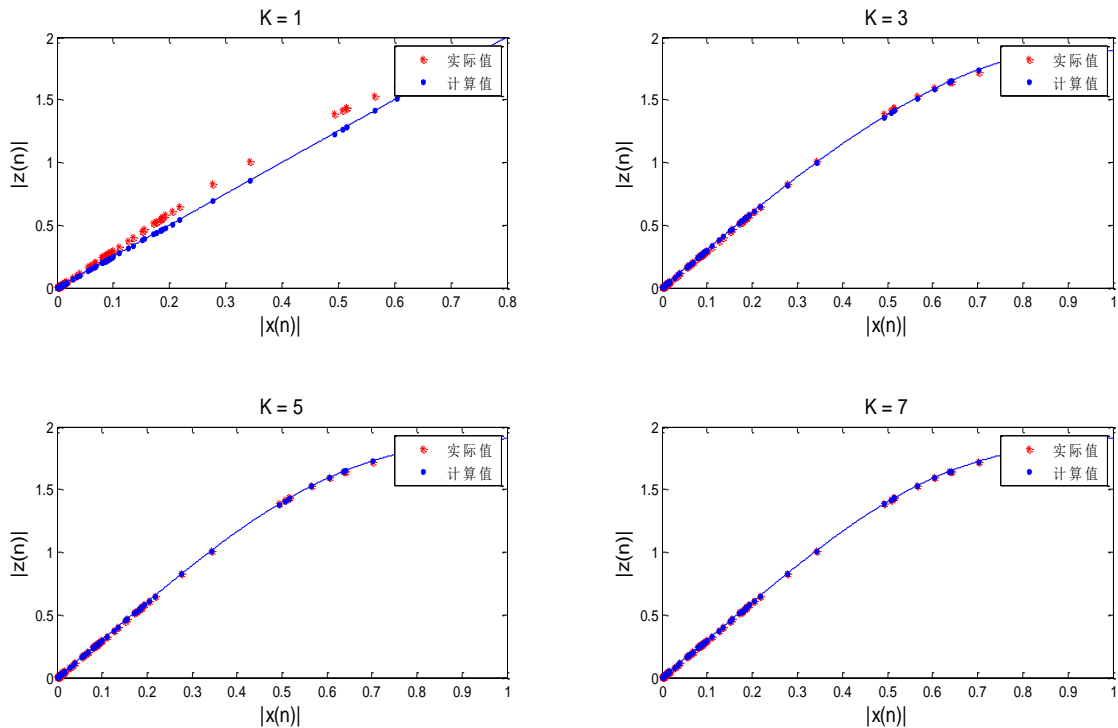


图6 实际分布情况的无记忆功放各阶多项式模型拟合效果图

表6 各阶实际分布正交基多项式模型及其准确度

阶次 K	数学模型表达式	NMSE
1	$y(n) = 2.50z(n)$	-16.34
3	$y(n) = 3.05z(n) + 0.01 z(n) z(n) - 1.17 z(n) ^2 z(n)$	-37.80
5	$y(n) = 2.97z(n) + 0.31 z(n) z(n) - 0.15 z(n) ^2 z(n) - 3.42 z(n) ^3 z(n) + 2.21 z(n) ^4 z(n)$	-53.44
7	$y(n) = 3.01z(n) - 0.28 z(n) z(n) + 1.94 z(n) ^2 z(n) - 4.60 z(n) ^3 z(n) - 3.20 z(n) ^4 z(n) + 8.84 z(n) ^5 z(n) - 3.80 z(n) ^6 z(n)$	-68.62

此计算结果与效果图表明，当 $K \leq 7$ 时，实际分布情况下的正交基多项式模型可用来表示无记忆功放的非线性特征，且随着阶次增加，准确度变高。

3.2.3 方案比较

从上述结果可以看出，当功放模型中的多项式阶数越高，模型的误差越小，为了验证测量的准确性，我们在第 2 部分定义了归一化均方误差（NMSE）如式（1）所示，其中 $z(n)$ 是功放的仿真输出， $\hat{z}(n)$ 是多项式模型的输出。图 7 给出了分别使用传统多项式模型、均匀分布情况下正交多项式模型与实际分布情况下正交多项式模型的归一化均方误差与多项式阶数 K 的关系曲线。

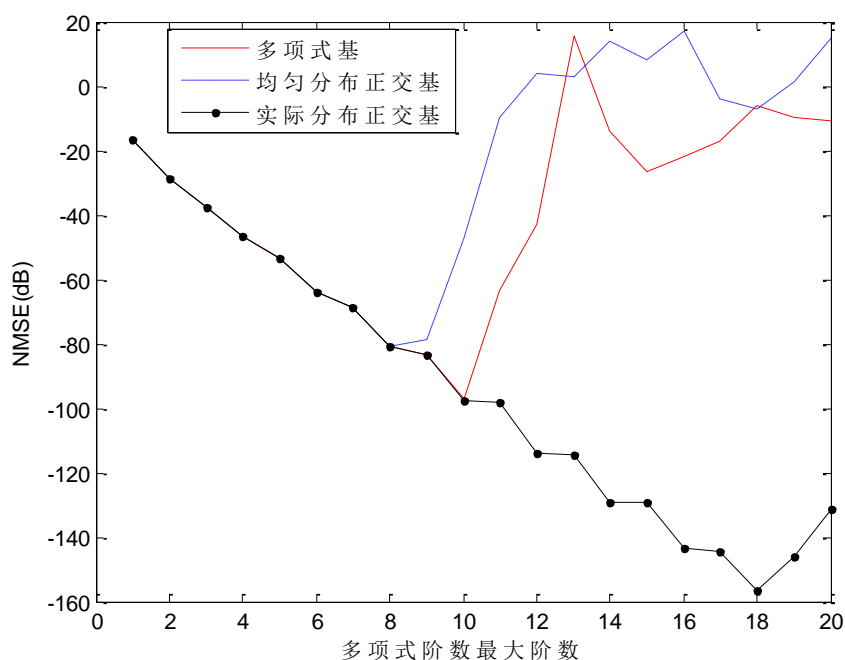


图7 归一化误差（NMSE）与多项式阶数的关系曲线

从图中可以看出，当多项式的最大阶数 K 小于 7 时，传统的多项式模型和正交多项式模型的误差基本相同，且 NMSE 均随 K 的增加而减小，即模型的准确度越来越好。当最大阶数 K 大于 7 时，对于实际分布情况的正交多项式模型来说，随着多项式最高阶数 K 的增加，NMSE 不断地减小，而对于传统的多项式模型和均匀分布情况的正交多项式模型来说，NMSE 反向增大。原因在于后两

者的 $\Phi^T \Phi$ 和 $\Psi^T \Psi$ 的条件数在 $K > 7$ 以后迅速变大，求逆稳定性变差，使得所求逆矩阵偏离真实值较远，所以 NMSE 的值变大，模型的准确度变差。所以当 K 大于 7 时，我们改进的按实际分布情况的正交多项式模型具有相当明显的优势。同时，利用正交多项式基所获得最小 NMSE 可达到 -157dB。远远超过多项式基及均匀分布下的正交多项式基。

然而在实际应用中，除了模型的精度之外，还需考虑到模型的泛化性能。随着模型阶数的增加，利用传统的最小二乘方法拟合得到的函数将会在训练样本上获得极小的误差。然而此时模型一方面会变得极为复杂，另一方面其泛化误差较大，即对训练样本之外的样本难以获得令人满意的精度。一种有效的度量在多项式拟合中过拟合情况的方法为模型中参数绝对值的最大值，我们定义模型的过拟合程度为 $O = \max(|\mathbf{a}|)$ 。下文中过拟合的衡量方法如无特殊说明皆与此处定义相同。

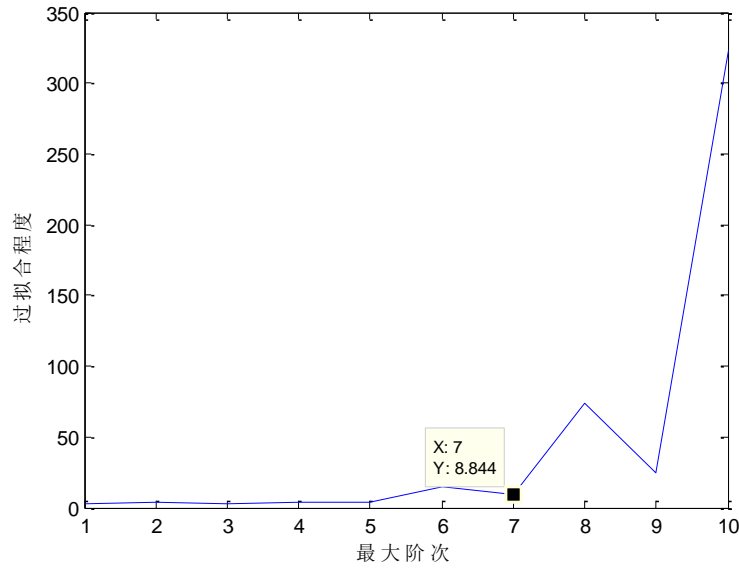


图 8 功放模型的过拟合程度与多项式最大系数之间的关系。

综合考虑模型的 NMSE 及 O ，我们选择实际分布情况的 7 阶正交多项式作为无记忆功放非线性特性的最优模型，如式 (15) 所示，其 NMSE=-68.63dB。

$$\begin{aligned}
 y(n) = & 3.01z(n) - 0.28|z(n)|z(n) + 1.94|z(n)|^2 z(n) - 4.60|z(n)|^3 z(n) \\
 & - 3.20|z(n)|^4 z(n) + 8.84|z(n)|^5 z(n) - 3.80|z(n)|^6 z(n)
 \end{aligned} \quad (15)$$

3.3 无记忆功放预失真模型的建立与求解

3.3.1 正交多项式预失真模型建立

预失真技术即在功放前加入一个非线性单元，称为预失真器，使其特性与 PA 特性互逆。因此信号预先人工产生了非线性失真，与放大器压缩非线性失真相抵消，实现整个系统的线性放大目的，原理如图 8 所示。

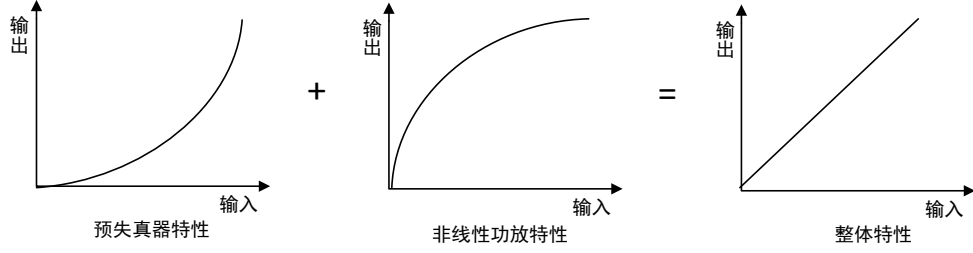


图9 预失真原理图

由文献[2]可知，预失真的实质为功放模型的求逆问题，理论上如果功放模型在信号包络区间是单调的，则其逆存在。由于直接求逆方法在实现上并不存在通用性，因此通常使用多项式逼近的方法进行预失真模型的构造，我们仍然采用实际输入信号幅度分布情况下的正交多项式模型，则如式（16）所示。

$$z(n) = \sum_{k=1}^K b_k x(n) |x(n)|^{k-1} = \sum_{k=1}^K b_k \psi_k(x(n)) \quad (16)$$

式中， $\psi_k(x(n))$ 为表4所示的正交基函数。

由文献[1]可知预失真处理技术通常采用的自适应算法，其控制结构有两种：直接学习结构和间接学习结构。如果直接对预失真器的参数进行自适应估计，需已知功放的模型，为了避免这个问题，我们选用图10所示的间接学习结构。

图中功放的线性放大倍数为 g ，功放的输出 $y(n)$ 经 g 倍衰减后作为预失真模块 A 的输入，实际预失真器的模型为其训练模块 A 的复制。定义误差信号 $e(n) = z(n) - \hat{z}(n)$ ，理想情况下，希望 $y(n) = gx(n)$ ，此时 $z(n) = \hat{z}(n)$ ， $e(n) = 0$ 。我们利用最小二乘算法，不断极小化误差 $e(n)$ ，通过训练使预失真训练模块的输出 $\hat{z}(n)$ 不断接近功放的输入 $z(n)$ 。当 $e(n)$ 最小时，即可得到训练模块的参数，从而得到预失真器模型的参数。

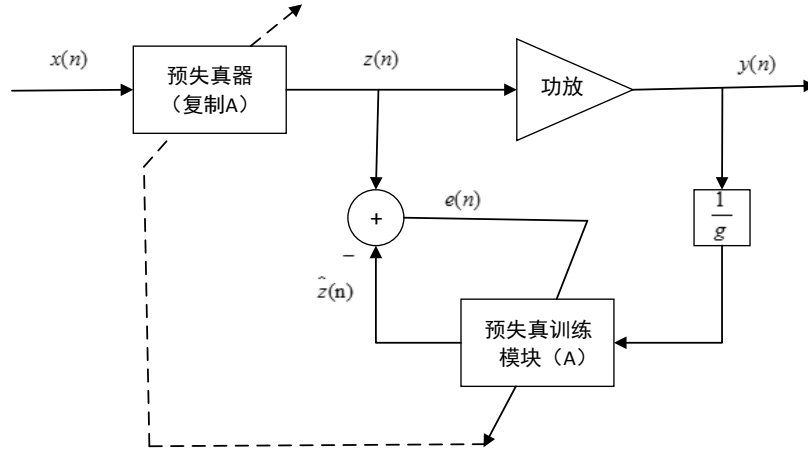


图10 数字预失真的非直接学习结构

由图10可定义预失真训练模块A的多项式模型为；

$$\hat{z}(n) = \sum_{k=1}^K b_k \frac{y(n)}{g} \left| \frac{y(n)}{g} \right|^{k-1} = \sum_{k=1}^K b_k \psi_k \left(\frac{y(n)}{g} \right) \quad (17)$$

我们定义输入和输出向量： $u(n) = \frac{y(n)}{g}$ ，

$\mathbf{u}=[u(1), \dots, u(n)]^T, \mathbf{z}=[z(1), \dots, z(n)]^T$ ，定义参数向量：

$\mathbf{b}=[b_1, b_2, \dots, b_k]^T$ ，再定义 $\mathbf{R}_k(u)=[R_k(u(1)), \dots, R_k(u(n))]^T$ ， $N \times K$ 矩阵

$\mathbf{R}=[\mathbf{R}_1(\mathbf{u}), \dots, \mathbf{R}_k(\mathbf{u})]$ ，则式 (16) 写成向量形式：

$$\mathbf{z} = \mathbf{R}\mathbf{b} \quad (18)$$

那么，多项式的系数可由最小二乘估计得到：

$$\hat{\mathbf{b}}_{LS} = [\hat{b}_1, \dots, \hat{b}_k]^T = (\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \mathbf{z} \quad (19)$$

其估计误差为：

$$e(n) = z(n) - \hat{z}(n) = z(n) - \sum_{k=1}^K b_k \frac{y(n)}{g} \left| \frac{y(n)}{g} \right|^{k-1} \quad (20)$$

由于 $e(n)$ 为实数，则定义其目标误差函数定义为：

$$J = \sum_{n=1}^N e(n) = \sum_{n=1}^N \left[z(n) - \sum_{k=1}^K b_k \frac{y(n)}{g} \left| \frac{y(n)}{g} \right|^{k-1} \right] \quad (21)$$

从上述计算理论可看出，要拟合得到多项式 (17) 的系数，必须首先确定线性放大倍数 g 。由本题所提供的资料分析得到， g 的取值受到预失真处理的“输出幅度限制”与功放的“输出功率最大化”所约束。

3.3.2 线性放大倍数 g 的确定

所谓的预失真处理的“输出幅度限制”是指功放的输入幅度需保持在一定范围，过大的值会导致饱和溢出。本题的建模中限定预失真处理的输出幅度不大于所给出的功放输入幅度最大值，由数据文件 1 可知，需要使预失真训练模块 A 的输出满足 $|\hat{z}(n)| \leq \max |x(n)| = 1.0553$ 。

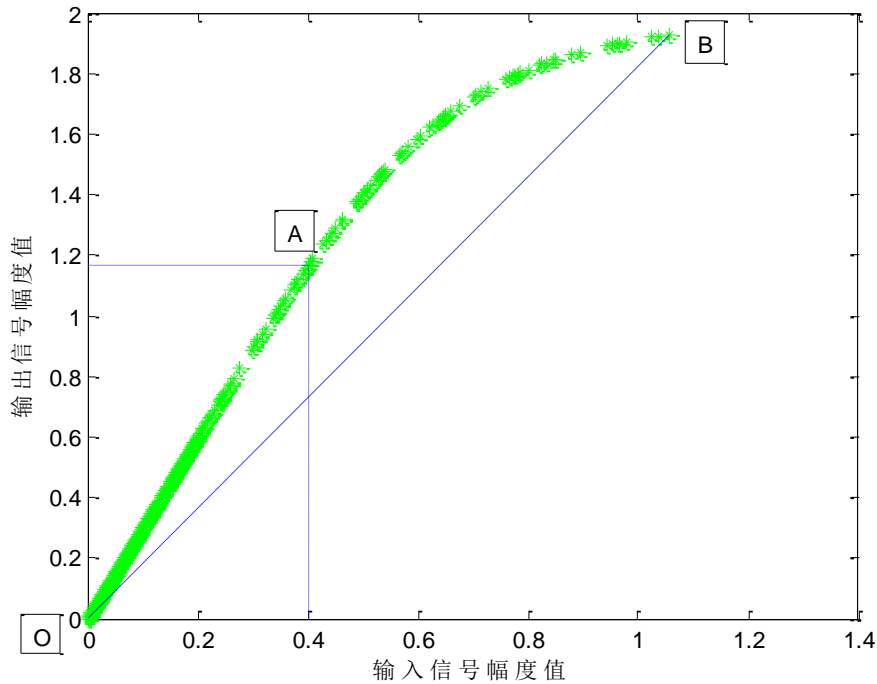


图 11 功放经预失真处理后整体的目标特性曲线

所谓的功放输出“功率最大化”是指预失真处理加载后，尽可能使功放的信

号平均输出功率最大，因此预失真处理后的输出幅度（等效为功放的输入幅度）需尽可能提高。图 11 是本题功放的输入输出幅度特性曲线，OA 段是其线性工作区域，之后发生非线性失真，进入饱和区。传统电路设计上，减轻非线性失真效应的方法是只取功放特性的线性段 OA（A(0.4005,1.165)），这样功放输出的最大幅度为 1.165，降低了输出功率，导致功放效率的降低。而线性化技术可以在不降低线性度的情况下提高功放的效率，为了提高功放的利用率，经过预失真处理后，功放的最大输出幅度应仍能达到原来的最大输出幅度 1.9274（B 点）。所以经预失真处理后理想的输入/输出特性曲线应为直线 OB。

综上所述，满足两个约束条件下的最大可能的线性幅度放大倍数为直线 OB 的斜率，即：

$$g = \frac{1.9274}{1.0553} = 1.8264 \quad (22)$$

3.3.3 正交多项式预失真模型求解

将所求的 g 代入式 (17)，根据数据文件 1，按前述求正交基的方法求得正交基函数如表 7 所示，用 MATLAB2012 编程实现最小二乘算法，拟合得到 $K=1,3,5,7$ 时的各阶预失真训练模块的正交多项式数学模型，即得到预失真器的模型，我们将公式 (20) 中的误差 e 按照公式 (1) 中的定义转换为 NMSE 并计算之，该 NMSE 见 12(a)，我们可以得到预失真模块对输入信号的重构最好时可达 -95dB。进一步考虑模型的过拟合程度，见图 12(b)。综合考虑，我们选取预失真模型的阶数 $K=5$ 。此时的正交基见表 6，所解得模型及参数见表 8。

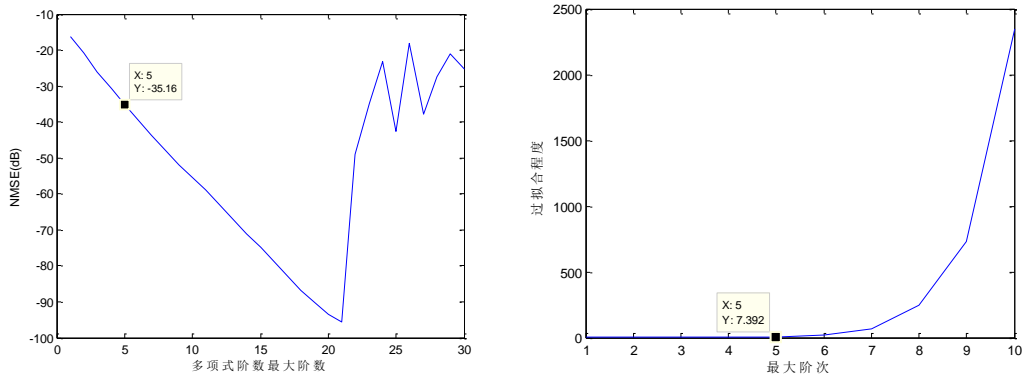


图 12 (a)NMSE 随多项式阶数增大的变化 (b)过拟合程度随多项式阶数增大的变化

表 7 数据文件 1 分布情况的正交基函数

$\psi_1(x) = -0.30 x ^4 x + 0.72 x ^3 x - 0.59 x ^2 x + 0.19 x x - 0.02x$
$\psi_2(x) = -0.48 x ^4 x + 0.39 x ^3 x + 0.54 x ^2 x - 0.56 x x + 0.10x$
$\psi_3(x) = 0.57 x ^4 x + 0.14 x ^3 x - 0.34 x ^2 x - 0.66 x x + 0.32x$
$\psi_4(x) = 0.46 x ^4 x + 0.38 x ^3 x + 0.25 x ^2 x - 0.02 x x - 0.76x$
$\psi_5(x) = 0.37 x ^4 x + 0.39 x ^3 x + 0.42 x ^2 x + 0.47 x x + 0.56x$

表8 计算得预失真正交多项式模型及其准确度

阶次 K	数学模型表达式	NMSE
5	$z(n) = 0.70x(n) - 1.23 x(n) x(n) + 4.9 x(n) ^2 x(n) - 7.39 x(n) ^3 x(n) + 3.85 x(n) ^4 x(n)$	-35.16

我们还需进一步评估所获得预失真模型矫正的效果。此时，定义理想线性放大器的放大增益为 g ，其由式（22）给定。我们计算理想输出值 $y = gx$ 与经过矫正后的输出之间的NMSE。其中 x 为输入复信号。取预失真器阶数为5，功放阶数为7，经计算得到经过预失真矫正后的功放输出与理想线性功放输出之间的NMSE为-37.08dB。

3.3.4 小结

综上，本预失真建模方法中：

$$\text{目标误差函数为： } e(n) = z(n) - \hat{z}(n) = z(n) - \sum_{k=1}^K b_k \frac{y(n)}{g} \left| \frac{y(n)}{g} \right|^{k-1}$$

$$\text{线性化后最大可能的幅度放大倍数为： } g = \frac{1.9274}{1.0553} = 1.8264$$

预失真模型为：

$$z(n) = 0.70x(n) - 1.23|x(n)|x(n) + 4.9|x(n)|^2 x(n) - 7.39|x(n)|^3 x(n) + 3.85|x(n)|^4 x(n)$$

此时预失真补偿后的结果与理想线性功放之间的误差NMSE=-37.08dB。

3.4 有记忆功放问题分析

所谓有记忆功放是指功放的某一时刻输出不仅与此时刻输入有关，而且与此前某一时间段的输入有关。对无记忆功放多项式模型离散后表示如下：

$$y(n) = \sum_{k=1}^K h_k z^k(n) = h_1 z(n) + h_2 z^2(n) + \dots + h_K z^K(n) \quad n = 1, 2, \dots, N \quad (23)$$

对式（23）增加记忆效应，可以写为：

$$\begin{aligned}
 y(n) = \sum_{k=1}^K \sum_{m=0}^M h_{km} z^k(n-m) = & h_{10} z(n) + h_{11} z(n-1) + \dots + h_{1M} z(n-M) \\
 & + h_{20} z^2(n) + h_{21} z^2(n-1) + \dots + h_{2M} z^2(n-M) \\
 & + \dots \quad \dots + \dots \\
 & + h_{K0} z^K(n) + h_{K1} z^K(n-1) + \dots + h_{KM} z^K(n-M)
 \end{aligned} \quad (24)$$

$n = 1, 2, \dots, N$

式中， M 为记忆深度， h_{km} 为系数。

由于常用复值输入-输出信号，可采用便于计算的“和记忆多项式”模型：

$$y(n) = \sum_{k=1}^K \sum_{m=0}^M h_{km} z(n-m) |z(n-m)|^{k-1} \quad n = 1, 2, \dots, N \quad (25)$$

其中， K 是多项式函数的最大阶数，阶次 K 越大，则模型的准确度越高。

因此，对于问题二，我们的求解思路如下：

(1) 基于对问题一的求解过程，我们发现基于正交基函数的多项式模型在表达功放的非线性特征时具有较大优势，于是，对有记忆功放的建模采用正交基“和记忆多项式”模型。首先利用数据文件2构建了符合实际分布的正交基函数。随后利用最小二乘拟合求解出数学模型及其NMSE。最后折中NMSE及过拟合程度，获得实际模型及对应NMSE。

(2) 对预失真模型的建立同样采用正交基多项式形式；因为模型具有记忆，难以直接分析出“输出幅度限制”和“功率最大化”两个约束条件下可能的线性化的最大幅度放大倍数；对于多种可能的放大倍数，采用自适应间接学习结构及最小二乘算法估算预失真模型的参数，找到最大的放大倍数及相应模型参数。最后与理想线性功放相比，得到预失真校正后的NMSE。

3.5 有记忆功放模型的建立与求解

3.5.1 有记忆功放模型建立

为了算法描述方便，我们首先定义 $z_{km}(n) = z(n-m)|z(n-m)|^{k-1}$ ， $\mathbf{z}_{km} = [z_{km}(1), z_{km}(2), \dots, z_{km}(n)]^T$ ，将式 (24) 改写成矩阵形式：

$$\mathbf{y} = \mathbf{Z}\mathbf{h} \quad (26)$$

式中， $\mathbf{y} = [y(1), \dots, y(n)]^T$ ，

$$\mathbf{Z} = [\mathbf{z}_{10}, \dots, \mathbf{z}_{K0}, \mathbf{z}_{11}, \dots, \mathbf{z}_{K1}, \dots, \mathbf{z}_{1M}, \dots, \mathbf{z}_{KM}]$$

$$\mathbf{h} = [h_{10}, \dots, h_{K0}, h_{11}, \dots, h_{K1}, \dots, h_{1M}, \dots, h_{KM}]^T$$

引入基于正交基函数的多项式模型，把式 (26) 改写成：

$$\mathbf{y} = \boldsymbol{\Psi}\mathbf{c} \quad (27)$$

式中， $\boldsymbol{\Psi}$ 是一组正交基，根据数据文件 2，按前述方法构造出的正交基函数，由于 $\boldsymbol{\Psi}$ 是 $K \times M$ 阶矩阵，若取 $K=7, M=5$ ，则正交基函数的个数为 35 个，由于数目较多，所有基函数的系数见附录。

在最小二乘条件下，系数的解析解为：

$$\hat{\mathbf{c}}_{LS} = (\boldsymbol{\Psi}^T \boldsymbol{\Psi})^{-1} \boldsymbol{\Psi}^T \mathbf{y} \quad (28)$$

那么，对于任意记忆深度 M 和最大阶次 K ，我们用 MATLAB2012 编程求解式 (28)，进行最小二乘拟合得到正交基多项式的系数。

3.5.2 有记忆功放模型求解

从图 13 中可以看出，当此正交基多项式模型的阶次 K 固定为 7 时，记忆深度 $M \geq 5$ 以后的 NMSE 几乎不变，表明此时模型的准确度不随记忆深度的增加而明显提高；当记忆深度 M 固定为 5 时，阶次 $K \geq 5$ 以后的 NMSE 也几乎不变，表明此时模型的准确度不随阶次的增加而明显提高；而图中右上角明显凸起的部分表明当阶次 K 和记忆深度 M 同时增大时不但不会减小 NMSE，反而使其迅速增大，则此时的模型准确度非常差。原因在于此时模型参数过多，难以获得精确解。在 $M > 5, K > 7$ 之后，可以看出模型 NMSE 基本不变。因此我们只对 $M \leq 5$ 和 $K \leq 7$ 的情况进行深入研究。考虑到三维图难以展示细节，我们将其映射至二维平面。如图 14(a) 是将记忆深度 M 和阶数 K 按 $(M, K) \rightarrow KN + M$ 进行映射后作出的二维平面图，其中 $M=5$ 为所需进一步研究的最大记忆深度。同时，依此画出过拟合程度随记忆深度 M 和阶数 K 的变化，如图 14(b)。

因此，综合考虑 NMSE 及过拟合程度，我们选取阶次为 $K=4$ ，记忆深度为 $M=5$ 的有记忆的正交基多项式模型较为合理，随着 M 和 K 的进一步增大，NMSE 的减小较为缓慢，然而过拟合程度却迅速增大。模型的具体参数见附件。

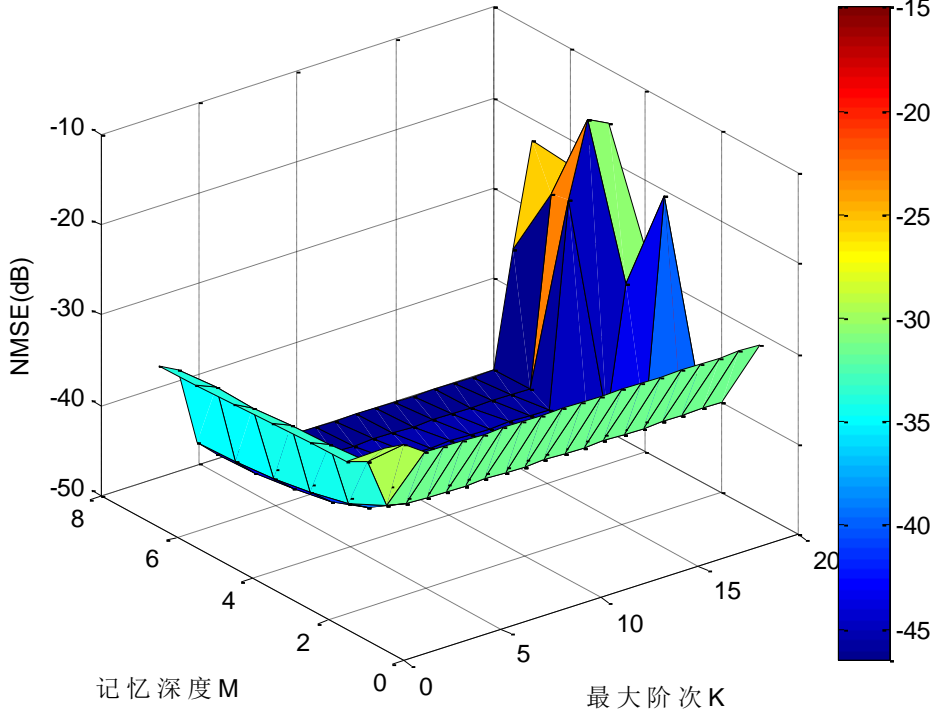


图13 NMSE与记忆深度及最大阶次的关系特性

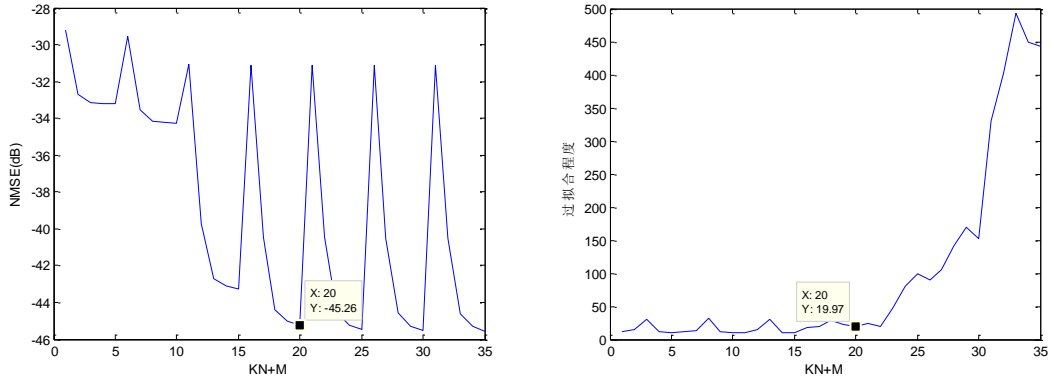


图14 (a) NMSE与记忆深度及最大阶次的关系 (b) 随参数组合增加过的拟合程度变化趋势

3.6 有记忆功放预失真模型的建立与求解

3.6.1 正交多项式预失真模型

与无记忆功放的预失真模型一样，对于有记忆功放的预失真模型可以采用记忆正交多项式模型来描述：

$$z(n) = \sum_{k=1}^K \sum_{m=0}^M p_{km} x(n-m) |x(n-m)|^{k-1} \quad n=1, 2, \dots, N \quad (29)$$

与无记忆功放的预失真模型一样，我们仍然采用自适应算法来估算有记忆功放的预失真模型，系统结构也采用间接学习结构。由于预失真器采用记忆多项式模型，在模型的阶数 K 和记忆深度 M 确定的情况下（在设计时， K 和 M 是可以根据实际情况改变的），采用最小二乘算法，通过极小化均方误差来拟合预失真训练模块记忆正交多项式模型的系数，即可得到相应的预失真器。本题采用最小二乘拟合法进行预失真器参数的估算。

参考前述图10，预失真训练模块的多项式模型表述为：

$$\hat{z}(n) = \sum_{k=1}^K \sum_{m=0}^M p_{km} \frac{y(n-m)}{g} \left| \frac{y(n-m)}{g} \right|^{k-1} \quad (30)$$

因有记忆的功放在工作时会产生 M 个时延，前 M 个输出无法估计，所以
 $n = M, \dots, N$

我们定义：

$$y_{km}(n) = y(n-m) \left| y(n-m) \right|^{k-1}, \mathbf{y}_{km} = [y_{km}(1), \dots, y_{km}(n)]^T$$

$$\mathbf{Y} = [\mathbf{y}_{10}, \dots, \mathbf{y}_{K0}, \mathbf{y}_{11}, \dots, \mathbf{y}_{K1}, \dots, \mathbf{y}_{1M}, \dots, \mathbf{y}_{KM}]$$

$$\hat{\mathbf{z}} = [\hat{z}(1), \dots, \hat{z}(n)]^T$$

$$\mathbf{p} = [p_{10}, \dots, p_{K0}, p_{11}, \dots, p_{K1}, \dots, p_{1M}, \dots, p_{KM}]^T$$

将式（30）改写成矩阵形式：

$$\hat{\mathbf{z}} = \mathbf{Y}\mathbf{p} \quad (31)$$

引入基于正交基函数的多项式模型，把式（26）改写成：

$$\hat{\mathbf{z}} = \boldsymbol{\Psi}\mathbf{d} \quad (32)$$

式中， $\boldsymbol{\Psi}$ 是一组正交基，根据数据文件2，按前述方法构造出的正交基函数，由于 $\boldsymbol{\Psi}$ 是 $K \times M$ 阶矩阵，若取 $K=4, M=5$ ，则正交基函数的个数为20个，由于数目较多，所有基函数的系数见附录，并作出该预失真模型的框图形式如图15。

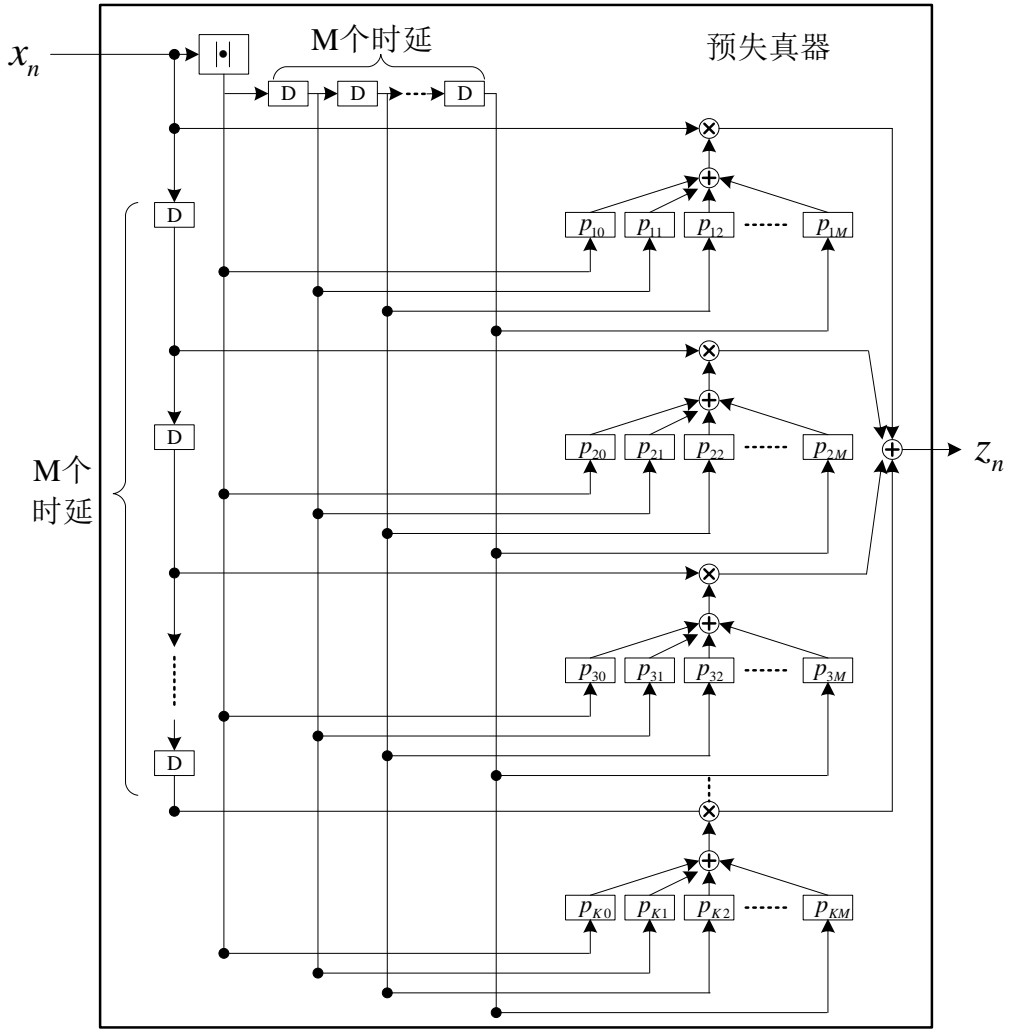


图 15 有记忆功放的预失真模型框图

模型中，多项式的系数仍由最小二乘估计得到：

$$\hat{d}_{LS} = [\hat{d}_1, \dots, \hat{d}_k]^T = (\Psi^T \Psi)^{-1} \Psi^T \hat{\mathbf{z}} \quad (33)$$

预失真训练模块的输出 \$\hat{z}(n)\$ 和预失真器的输出 \$z(n)\$ 之间的估计误差为：

$$\begin{aligned} e(n) &= z(n) - \hat{z}(n) \\ &= z(n) - \sum_{k=1}^K \sum_{m=0}^M p_{km} \frac{y(n-m)}{g} \left| \frac{y(n-m)}{g} \right|^{k-1} \quad n = M, \dots, N \end{aligned} \quad (34)$$

引入：

$$\begin{aligned} &\sum_{n=M}^N e^*(n) e(n) \\ &= \sum_{n=M}^N \left[z(n) - \sum_{k=1}^K \sum_{m=0}^M p_{km} \frac{y(n-m)}{g} \left| \frac{y(n-m)}{g} \right|^{k-1} \right]^* \left[z(n) - \sum_{k=1}^K \sum_{m=0}^M p_{km} \frac{y(n-m)}{g} \left| \frac{y(n-m)}{g} \right|^{k-1} \right] \end{aligned} \quad (35)$$

定义误差向量 \$\mathbf{e}\$ 的平方和为目标误差函数，即：

$$J = \|\mathbf{e}\|^2 = \mathbf{e}^* \mathbf{e} = \sum_{n=M}^N e^*(n)e(n) \quad (36)$$

为求系数，需极小化误差目标函数 J 。

从上述计算理论可看出，要拟合得到多项式系数，必须首先确定线性放大倍数 g 。由本题所提供的资料分析得到， g 的取值受到预失真处理的“输出幅度限制”与功放的“输出功率最大化”所约束。

3.6.2 线性放大倍数 g 的确定

本题的建模中限定预失真处理的输出幅度不大于所给出的功放输入幅度最大值，由数据文件 2 可知，需要使预失真训练模块 A 的输出满足 $|\hat{z}(n)| \leq \max |x(n)| = 0.7079$ 。

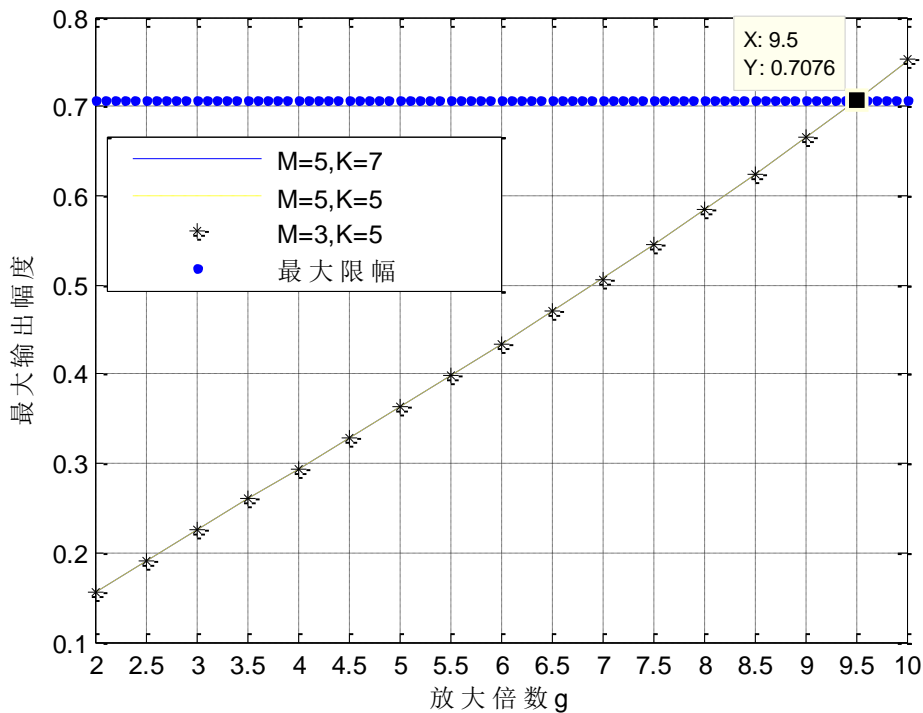


图 16 最大输出幅度与放大倍数 g 的关系曲线

满足功放输出“功率最大化”，需要满足预失真处理后的输出幅度（等效为功放的输入幅度）需尽可能提高。由于预失真模型的参数估计受放大倍数 g 的影响，不同的 g 会训练得到不同的预失真模型，则对应相同的输入样本，不同的模型得到的最大输出幅度不同，因此我们需找到最大输出幅度与放大倍数 g 之间关系。我们改变 g 的取值，同时改变记忆深度 M 及阶数 K 的取值。在不同参数配置下的预失真器输出的最大幅值如图 16 所示。从图中我们可看出，随着 g 的变化， M 和 K 的变化对最大输出幅值影响较小。另外，当放大倍数 g 取 9.5 时对应的最大输出幅度为 0.7076（小于功放的输入幅度最大值）。进一步增大 g 则会明显导致预失真器输出信号幅值大于限幅。因此，同时考虑上述的输出幅度限制条件和功率最大化，则有记忆功放加预失真矫正线性化后最大可能的放大倍数取 $g=9.5$ 。

3.6.3 有记忆功放预失真模型求解

将所求的 $g=9.5$ 代入式 (30)，根据数据文件 2，按前述求正交基的方法求得

正交基函数，用 MATLAB2012 编程实现最小二乘算法，可得到重构 NMSE 随记忆深度 M 及最大阶次 K 的变化如图 17(a)所示。同时得到过拟合程度的变化如图 17(b)所示。与之前分析过程类似。对于有记忆功放的预失真器取记忆深度 $M=5$ 和最大阶次 $K=5$ 的正交多项式模型，其具体表达式见附件。

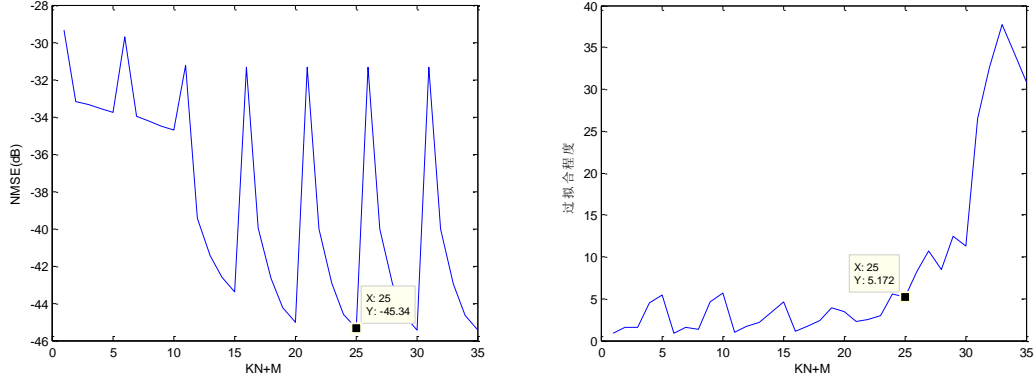


图17 (a) NMSE与记忆深度及最大阶次的关系 (b) 随参数组合增加的过拟合程度变化趋势

上述计算结果可知，预失真器的正交多项式模型随着阶次增加，准确度越来越好，考虑到阶次的增高会增加工程实现的复杂度，因此本预失真器模型选择 $K=7$ 阶正交多项式模型。

进一步评价总体功放模型，考虑放大倍数 $g=9.5$ 的理想线性功放。利用题目中给定的输入 pa_in_memory ，通过该理想线性功放的输出作为理想输出 z ，评价其和经过预失真和功放的输出 \hat{z} 之间的 NMSE。对于放大倍数 $g=9.5$ ，记忆深度 $M=5$ ，最大阶次 $K=5$ 的预失真器，经过记忆深度 $M=5$ ，最大阶次 $K=4$ 的功放模型，经计算得到矫正后与理想输出之间的 NMSE 为 -43.40dB。

3.6.4 小结

综上，本预失真建模方法中：

目标误差函数为：

$$J = \|\mathbf{e}\|^2 = \mathbf{e}^* \mathbf{e} = \sum_{n=M}^N e^*(n)e(n)$$

$$= \sum_{n=M}^N \left[z(n) - \sum_{k=1}^K \sum_{m=0}^M p_{km} \frac{y(n-m)}{g} \left| \frac{y(n-m)}{g} \right|^{k-1} \right]^* \left[z(n) - \sum_{k=1}^K \sum_{m=0}^M p_{km} \frac{y(n-m)}{g} \left| \frac{y(n-m)}{g} \right|^{k-1} \right]$$

线性化后最大可能的幅度放大倍数为： $g = 9.5$

预失真模型为： 见附件

其 NMSE = -43.40dB

3.7 相邻功率信道比及功率谱密度

3.7.1 问题分析

从本题题目所给材料可获取以下信息：

(1) 相邻信道功率比 (Adjacent Channel Power Ratio, ACPR) 是表示信道的带外失真的参数，衡量由于非线性效应所产生的新频率分量对邻道信号的影响程度。其定义为

$$ACPR = 10 \log_{10} \frac{\int_{f_2}^{f_3} s(f) df}{\int_{f_1}^{f_2} s(f) df} \quad (37)$$

其中 $s(f)$ 为信号的功率谱密度函数， $[f_1, f_2]$ 为传输信道， $[f_2, f_3]$ 为相邻信道。

(2) 功率谱密度 $s(f)$ 的计算可通过对信号的自相关函数进行 Fourier 变换计算。

(3) 假设数据 2 的采样频率 $F_s = 30.72 \times 12 \text{ MHz}$ ，传输信道按照 20MHz 来算，邻信道也是 20MHz。

通过查阅相关资料，此题我们的求解思路如下：

(1) 利用 MATLAB 编程求解三类信号（输入信号、无预失真补偿的功率放大器输出信号、采用预失真补偿的功率放大器输出信号）的 $s(f)$ ，其中取 $2^n=256$ 的 FFT。

(2) 对三类信号的 ACPR 的求解，需要已知各自的 $s(f)$ ，传输信道 $[f_1, f_2]$ ，相邻信道 $[f_2, f_3]$ ，其中， f_1, f_2, f_3 可由第一步中求得的各自的功率谱密度图中看出。

3.7.2 三类信号的功率谱密度求解

由于 MATLAB 自带的求解功率谱密度的算法即为傅里叶变换法，即先由信号序列 $x(n)$ 估计出自相关函数 $R(n)$ ，然后 $R(n)$ 进行傅立叶变换，便得到 $x(n)$ 的功率谱估计。因此直接利用 MATLAB 求解，取 256 点 FFT，对输入信号序列 $x(n)$ 和无预失真补偿的功率放大器输出信号序列 $y(n)$ ，直接由数据文件 2 获得，对采用预失真补偿的功率放大器输出信号序列 $\hat{y}(n)$ ，则是如图所示的系统的输出，该系统中预失真处理器模型及 PA 模型的参数见附录，其中 $g=9.5$ ， $K=5$ ， $M=5$ 。

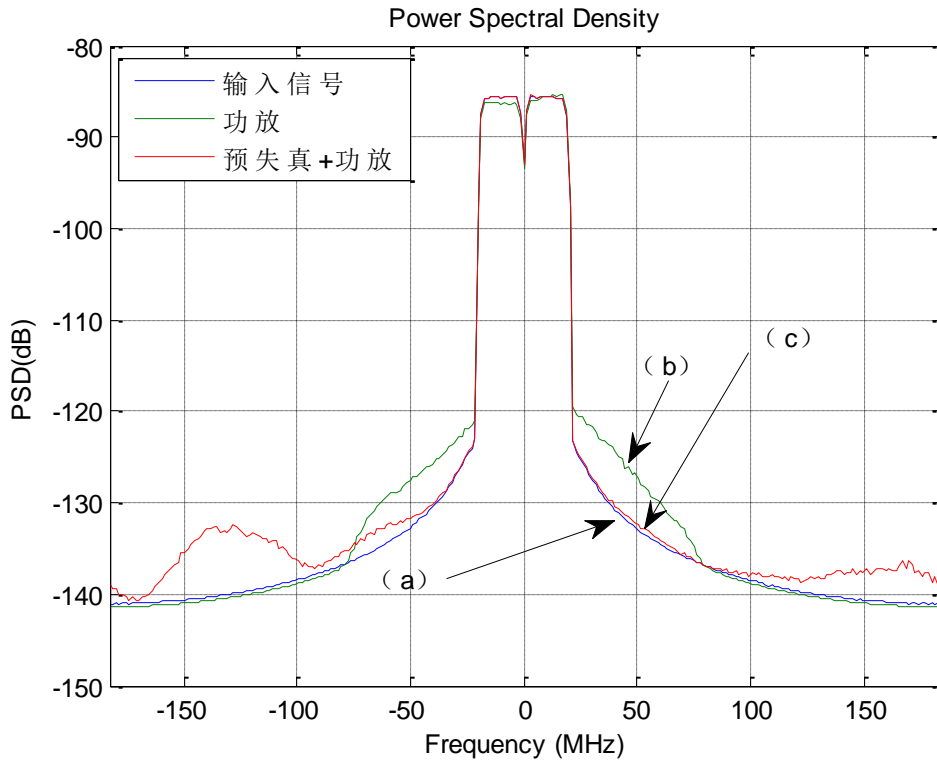


图18 功放预失真补偿前后的功率谱密度比较

最终求得三种信号的功率谱密度如图18所示。其中(a)表示输入信号的功率谱

密度, (b)表示仅通过功率放大器时输出信号的功率谱密度, (c)表示输入信号通过预失真器及功率放大器之后输出的功率谱密度。为了方便对比, 图中(b)和(c)对输入信号的放大作用已被消除。

3.7.3 结果分析

由图18可知, 在传输信道[0,20MHz]和相邻信道[20MHz-40MHz]上, 预失真处理后的功率谱密度近似相等, 可见预失真补偿后的输出信号明显改善。表9是数据2经预失真处理前后的ACPR对比, 可以发现, 经预失真器处理后信号可得到约-15dB的补偿, g 的取值对ACPR值的影响可以忽略。

表 9 预失真处理前后的 ACPR 对比

ACPR 值(MPre = 5, KPre=5, MPA=5, KPA=4)				
原始信号	功放	预失真+功放 $g=9.5$	预失真+功放 $g=9$	预失真+功放 $g=8.5$
-78.15	-39.38	-54.34	-54.28	-54.14

其中 MPre 为预失真器的记忆深度, KPre 为预失真器的最大阶数, MPA 为功放的记忆深度, KPA 为功放的最大阶数。

4 模型改进

本文为防止过拟合,采用了较为主观的参数选择方法。另一种常见的防止过拟合的方法是在最小二乘法的优化目标后加入防止过拟合的正则项,然而该方法的正则项的系数也由人来主观设置。

在最小二乘法的正则项中还可以加入其他约束条件。例如,如果考虑工程实现时用尽可能少的元器件完成预失真器,则可加入稀疏约束条件。此时最小二乘方法则成为LASSO方法。由于时间有限,本文并未实现LASSO方法。然而LASSO方法在实际应用中应得到重视。

对于一般的曲线拟合问题,神经网络大多可取得较好的拟合效果。然而对于预失真器的设计问题,国际学者极少利用神经网络进行参数拟合,但本方法不失为一种值得尝试的方法。

参考文献

- [1] J. Kim and K. Konstantinou, Digital predistortion of wideband signals based on power amplifier model with memory, Electron. Lett., vol. 37,no. 23,pp.1417-1418, Nov. 2001.
- [2] Ding L, Zhou G T, Morgan D R, et al. A robust digital baseband predistorter constructed using memory polynomials[J]. Communications, IEEE Transactions on, 2004, 52(1): 159-165.
- [3] Raich R, Qian H, Zhou G T. Orthogonal polynomials for power amplifier modeling and predistorter design[J]. Vehicular Technology, IEEE Transactions on, 2004, 53(5): 1468-1479.
- [4]何子述,夏威等. 现代数字信号处理及其应用.北京:清华大学出版社,2009.
- [5]龚耀寰. 自适应滤波—时域自适应滤波和智能天线(第二版).北京:电子工业出版社,2003:308-333

附录

1. 有记忆功放模型中采用的正交基系数：
见附件“pa_basis.csv”
2. 有记忆功放模型得到的和记忆公式的系数：
见附件“pa_param.csv”
3. 有记忆功放的预失真模型计算时采用的正交基系数：
见附件“pre_basis.csv”
4. 有记忆功放的预失真模型算得的和记忆公式的系数：
见附件“pre_param.csv”