

# ARMAX 模型两种自适应 $k$ 步超前 预报形式的比较

姜启源 孙义军

(清华大学应用数学系, 北京 100084)

## 摘 要

ARMAX 模型的自适应  $k$  步超前预报通常有单步预报和多步预报两种形式, 前者依赖于若干个过去时刻固定步数 ( $k$  步) 的预报, 后者则用到一系列固定时刻不同步数 ( $k-1, k-2, \dots$ ) 的预报. 本文证明二者是完全等价的, 并且从应用的角度和模拟试验的结果对它们进行了比较.

**关键词:** ARMAX 模型,  $k$  步超前预报, 单步预报形式, 多步预报形式.

## 一、前 言

众所周知, 相当广泛的一类工业生产过程及许多领域的动态系统可以用 ARMAX (带外部变量的自回归滑动平均) 模型来描述. 根据这个模型对输出量作  $k$  步超前预报, 是指基于直到时刻  $t$  的输出及各个时刻已知的或设定的输入, 对时刻  $t+k$  的输出作出估计. 一般说来, 作预报时模型的参数是给定不变的, 而自适应预报则是在根据每个时刻得到的输入、输出量适时辨识和修正模型参数的同时, 对输出进行估计.

ARMAX 模型的  $k$  步超前预报通常有单步预报和多步预报两种形式, 它们的导出和相应的算法已有一些讨论<sup>[1-4]</sup>. 本文在简要回顾这两种形式之后证明二者是完全等价的, 并且主要从应用的角度及模拟试验的结果对它们进行比较.

本文讨论的是单输入单输出模型, 但其结果可直接推广至多输入单输出模型, 至于多输入多输出模型将另文叙述.

## 二、两种预报形式

ARMAX 模型可用如下方程表示:

$$\begin{cases} A(z^{-1})y(t) = B(z^{-1})u(t) + C(z^{-1})e(t), \\ A(z^{-1}) = 1 + a_1z^{-1} + \dots + a_nz^{-n}, \\ B(z^{-1}) = b_1z^{-1} + \dots + b_mz^{-m}, \\ C(z^{-1}) = 1 + c_1z^{-1} + \dots + c_pz^{-p}. \end{cases} \quad (1)$$

其中  $y(t)$ ,  $u(t)$  分别是时刻  $t$  的输出和输入;  $e(t)$  是零均值白噪声序列, 方差为  $\sigma^2$ ;  $z^{-1}$  是单位延迟算子, 即  $z^{-1}y(t) = y(t-1)$ ; 多项式  $A(z^{-1})$  和  $C(z^{-1})$  互素, 且根均在单位圆内.

实际上应用的输出预报是最小方差预报. 记时刻  $t$  的  $k$  步超前预报为  $\hat{y}(t+k|t)$ , 它使准则函数——预报误差方差

$$J = E[y(t+k) - \hat{y}(t+k|t)]^2 \quad (2)$$

达到极小.

假设模型(1)中的诸参数  $A(z^{-1})$ ,  $B(z^{-1})$ ,  $C(z^{-1})$  已知或已由输入输出数据辨识得到, 下面给出  $k$  步超前预报的两种形式:

### 1. 单步预报<sup>[1,3]</sup>

$$\hat{y}(t+k|t) = \frac{G(z^{-1})}{C(z^{-1})} y(t) + \frac{z^k B(z^{-1}) F(z^{-1})}{C(z^{-1})} u(t), \quad (3)$$

其中  $G(z^{-1})$  和  $F(z^{-1})$  由下式确定:

$$\begin{cases} C(z^{-1}) = A(z^{-1})F(z^{-1}) + z^{-k}G(z^{-1}), \\ F(z^{-1}) = 1 + f_1 z^{-1} + \cdots + f_{k-1} z^{-(k-1)}, \\ G(z^{-1}) = g_0 + g_1 z^{-1} + \cdots + g_r z^{-r}, \quad r = \max(n-1, p-k). \end{cases} \quad (4)$$

(3), (4)式的展开形式为

$$\begin{aligned} \hat{y}(t+k|t) = & - \sum_{i=1}^p c_i \hat{y}(t+k-i|t-i) + \sum_{i=0}^r g_i y(t-i) \\ & + \sum_{i=1}^{m+k-1} h_i u(t+k-i), \end{aligned} \quad (5)$$

其中  $h_i$  是下列多项式的系数:

$$H(z^{-1}) = B(z^{-1})F(z^{-1}) = \sum_{i=1}^{m+k-1} h_i z^{-i}. \quad (6)$$

由(5)式可知, 为了计算  $\hat{y}(t+k|t)$  需要过去时刻  $t-1, t-2, \cdots, t-p$  的固定步数为  $k$  的预报, 所以称为单步预报. 对于准则函数(2)式这种形式的预报达到的最小方差是

$$E[y(t+k) - \hat{y}(t+k|t)]^2 = \sigma^2(1 + f_1^2 + \cdots + f_{k-1}^2). \quad (7)$$

### 2. 多步预报<sup>[2-4]</sup>

为了与单步预报  $\hat{y}(t+k|t)$  相区别, 将多步预报记作  $y(t+k|t)$ , 有

$$\begin{aligned} y(t+k|t) = & - \sum_{i=1}^n a_i y(t+k-i|t) + \sum_{i=1}^m b_i u(t+k-i) \\ & + \sum_{i=k}^p c_i e(t+k-i), \end{aligned} \quad (8)$$

式中设  $k \leq p$ . 当  $k > p$  时右端最后一项消失, 并且注意到

$$y(t+j|t) = y(t+j), \quad j \leq 0, \quad (9)$$

和  $e(t)$  等于 1 步预报误差, 即

$$e(t) = y(t) - y(t|t-1) \quad (10)$$

与单步预报不同, (8)式右端包含了固定时刻  $t$  的不同步数  $(k-1, k-2, \dots, 1)$  的预报, 所以称为多步预报. 这种预报误差的方差可由模型(1)隐含的 ARMA 模型  $A(z^{-1})y(t) = C(z^{-1})e(t)$  的传递形式

$$\begin{cases} y(t) = \tilde{F}(z^{-1})e(t) \\ \tilde{F}(z^{-1}) = \frac{C(z^{-1})}{A(z^{-1})} = 1 + \sum_{i=1}^{\infty} \tilde{f}_i z^{-i} \end{cases} \quad (11)$$

给出, 有

$$E[y(t+k) - y(t+k|t)]^2 = \sigma^2(1 + \tilde{f}_1^2 + \dots + \tilde{f}_{k-1}^2), \quad (12)$$

将(11)式与(4)式比较, 显然  $f_i = \tilde{f}_i$  ( $i = 1, 2, \dots, k-1$ ). 故由(7), (12)式得到

$$E[y(t+k) - \hat{y}(t+k|t)]^2 = E[y(t+k) - y(t+k|t)]^2, \quad (13)$$

即由(5)式定义的单步预报  $\hat{y}(t+k|t)$  和由(8)式定义的多步预报  $y(t+k|t)$  都是最小方差预报.

### 三、两种预报形式的等价性

已经看到, 单步预报  $\hat{y}(t+k|t)$  和多步预报  $y(t+k|t)$  有相同的预报误差最小方差, 下面证明它们是完全等价的, 即对于任意的  $k$ , 有

$$\hat{y}(t+k|t) = y(t+k|t). \quad (14)$$

证明. 采用归纳法.

首先, 当  $k=1$  时, 由(8)式并注意到(9)式,

$$\begin{aligned} y(t+1|t) = & - \sum_{i=1}^n a_i y(t-i+1) + \sum_{i=1}^m b_i u(t-i+1) \\ & + \sum_{i=1}^p c_i e(t-i+1), \end{aligned}$$

即

$$y(t+1|t) = z[(c(z^{-1}) - 1)e(t) - (A(z^{-1}) - 1)y(t)] + zB(z^{-1})u(t).$$

将(1)式中  $y(t)$  代入, 并注意到  $k=1$  时有  $C(z^{-1}) = A(z^{-1}) + z^{-1}G(z)$ , 则

$$y(t+1|t) = \frac{G(z^{-1})}{A(z^{-1})} e(t) + \frac{zB(z^{-1})}{A(z^{-1})} u(t). \quad (15)$$

而由(3)式得

$$\hat{y}(t+1|t) = \frac{G(z^{-1})}{C(z^{-1})} y(t) + \frac{B(z^{-1})}{C(z^{-1})} u(t+1).$$

同样地利用(1)式中的  $y(t)$  和  $C(z^{-1}) = A(z^{-1}) + z^{-1}G(z)$ , 有

$$\hat{y}(t+1|t) = \frac{G(z^{-1})}{A(z^{-1})} e(t) + \frac{zB(z^{-1})}{A(z^{-1})} u(t). \quad (16)$$

比较(15)与(16)式, 得

$$\hat{y}(t+1|t) = y(t+1|t). \quad (17)$$

其次, 假设

$$\hat{y}(t+s|t) = y(t+s|t), \quad s = 1, 2, \dots, k-1 \quad (18)$$

成立, 证明  $s = k$ , 即(14)式成立.

为了明确起见, 对于  $s$  步的单步预报将(3), (4)式重新写作

$$\hat{y}(t+s|t) = \frac{G_s(z^{-1})}{C(z^{-1})} y(t) + \frac{z^s B(z^{-1}) F_s(z^{-1})}{C(z^{-1})} u(t), \quad (19)$$

$$\begin{cases} C(z^{-1}) = A(z^{-1}) F_s(z^{-1}) + z^{-s} G_s(z^{-1}), \\ F_s(z^{-1}) = 1 + f_1 z^{-1} + \dots + f_{s-1} z^{-(s-1)}, \\ G_s(z^{-1}) = g_0^{(s)} + g_1^{(s)} z^{-1} + \dots + g_r^{(s)} z^{-r}, \quad r = n-1. \end{cases} \quad (20)$$

在(19)式中令  $s = k$  并将模型(1)代入得

$$\hat{y}(t+k|t) = \frac{G_k(z^{-1})}{A(z^{-1})} e(t) + \frac{z^k B(z^{-1})}{A(z^{-1})} u(t). \quad (21)$$

对于多步预报, 不妨设  $k \leq p \leq n$  (其余情况更为简单), 于是由(8), (9)式给出

$$\begin{aligned} y(t+k|t) = & - \sum_{i=1}^{k-1} a_i y(t+k-i|t) - \sum_{i=k}^n a_i z^{k-i} y(t) \\ & + \sum_{i=1}^m b_i z^{k-i} u(t) + \sum_{i=k}^p c_i z^{k-i} e(t). \end{aligned}$$

将(18), (19), (20)式及模型(1)代入上式右端整理后可得

$$\begin{aligned} y(t+k|t) = & z^k \left[ \frac{z^{-k} G_k(z^{-1})}{A(z^{-1})} + \sum_{i=1}^{k-1} a_i z^{-i} F_{k-i}(z^{-1}) \right. \\ & \left. + F_k(z^{-1}) - 1 - \sum_{i=1}^{k-1} c_i z^{-i} \right] e(t) + \frac{z^k B(z^{-1})}{A(z^{-1})} u(t). \end{aligned} \quad (22)$$

其中  $e(t)$  的部分系数

$$\sum_{i=1}^{k-1} a_i z^{-i} F_{k-i}(z^{-1}) + F_k(z^{-1}) - 1 - \sum_{i=1}^{k-1} c_i z^{-i} \quad (23)$$

恰是

$$C(z^{-1}) - A(z^{-1}) F_k(z^{-1}) = z^{-k} G_k(z^{-1}) \quad (24)$$

左端多项式中含  $1, z^{-1}, \dots, z^{-(k-1)}$  的前  $k$  项, 而(24)式右端  $z^{-1}$  的最低阶项为  $z^{-k}$ , 故(23)式为零. 于是由(22)式给出

$$\hat{y}(t+k|t) = \frac{G_k(z^{-1})}{A(z^{-1})} e(t) + \frac{z^k B(z^{-1})}{A(z^{-1})} u(t), \quad (25)$$

将(25)式与(21)式比较立即得到(14)式.

#### 四、两种预报形式的比较

实际上究竟采用哪种预报形式应视问题的性质和需要而定. 单步预报  $\hat{y}(t+k|t)$  包含着过去时刻的固定步数 ( $k$  步) 的一系列预报  $\hat{y}(t+k-i|t-i)$ , 根据每一时刻观

1) 为简单起见设  $p \leq n$ , 于是  $r = \max(n-1, p-s) = n-1$ . 其余情况可类似地讨论.

测到的输入输出数据, 只需递推地计算一个也只能得到一个  $k$  步超前预报 (如果  $k$  取值不同, 则需重新计算), 所以适用于控制输入固定地滞后于输出  $k$  步的一类工业过程的自适应预报, 以及基于预报的控制. 而多步预报  $y(t+k|t)$  含有同一时刻  $t$  的从 1 步直到  $k-1$  步的一系列预报  $y(t+k-i|t)$ , 根据每一时刻观测到的输入输出数据, 需要递推地计算  $k$  个也能够得到  $k$  个 (从 1 步到  $k$  步的) 超前预报, 有些工业过程, 特别是一些属于经济、社会、水文、气象等领域的问题, 常常需要这样的一系列不同步数的预报.

从计算角度看, 当模型(1)的参数  $A(z^{-1})$ ,  $B(z^{-1})$ ,  $C(z^{-1})$  未知时有两种自适应预报方法, 即间接法和直接法. 前者先由输入输出数据辨识模型参数 (可以在线进行递推辨识), 再计算输出预报; 后者则直接由输入输出数据给出预报值. 下面分别对单步预报和多步预报就这两种方法作简要讨论.

采用间接法时, 对单步预报应先用 RELS (递推增广最小二乘) 或者 RPE (递推预报误差), RML (递推极大似然) 等方法辨识  $A(z^{-1})$ ,  $B(z^{-1})$ ,  $C(z^{-1})$ , 并用(4)式确定  $F(z^{-1})$ ,  $G(z^{-1})$ , 再由(3)式得到预报  $\hat{y}(t+k|t)$ ; 对多步预报则只需辨识出  $A(z^{-1})$ ,  $B(z^{-1})$ ,  $C(z^{-1})$ , 即可用(8)式算出预报  $y(t+k|t)$  (但先要计算  $k-1$  个预报  $y(t+s|t)$ ,  $s=1, 2, \dots, k-1$ ). 可以看出, 当  $k$  较小时 (如  $k=1, 2, 3$ ) 多步预报的计算量较小.

为了比较预报精度做如下模拟试验:

对于模型(1), 参数真值取为

$$A(z^{-1}) = 1 - 1.5z^{-1} + 0.7z^{-2},$$

$$B(z^{-1}) = z^{-1} + 0.5z^{-2},$$

$$C(z^{-1}) = 1 + 0.7z^{-1} + 0.2z^{-2}.$$

输入  $u(t)$  取幅度为 1 的 4 阶  $M$  序列,  $e(t)$  取零均值正态白噪声, 均方差  $\sigma = 0.1$ . 对长度为 800 的仿真数据先用 RELS 方法辨识参数, 再自  $t = 750$  开始分别用单步和多步预报形式进行  $k=1$ ,  $k=2$  的自适应预报, 结果如表 1 所示 (表中只列出 50 组预报数据中的最后 20 个结果, 预报误差方差为 50 组数据的平均值).

模拟试验表明, 两种预报的结果非常接近, 预报误差方差几乎一样, 且与(7)式给出的理论值相差不大. 虽然一般说来,  $C(z^{-1})$  的辨识精度不如  $A(z^{-1})$ ,  $B(z^{-1})$ , 似乎影响单步预报(3)式的精度, 但是由于(3)式可以转换为(21)式, 与多步预报(8)式比较, 可知参数辨识的结果对预报的影响是相近的.

采用直接法时, 对于单步预报可定义参数集和数据集分别为

$$\begin{cases} \theta = (c_1, c_2, \dots, c_p; g_0, g_1, \dots, g_r; h_1, h_2, \dots, h_{m+k-1})^T, \\ \varphi(t, \theta) = (-\hat{y}(t+k-1|t-1, \theta), -\hat{y}(t+k-2|t-2, \theta), \dots, \\ \quad -\hat{y}(t+k-p|t-p, \theta); y(t), y(t-1), \dots, y(t-r); \\ \quad u(t+k-1), u(t+k-2), \dots, u(t-m+1))^T, \end{cases} \quad (26)$$

则预报公式(5)可表示为

$$\hat{y}(t+k|t, \theta) = \varphi^T(t, \theta) \hat{\theta}(t), \quad (27)$$

其中  $\hat{\theta}(t) = \hat{\theta}$  是时变参数在时刻  $t$  的估计值, 用 RPE 方法可由如下一套递推公式计算:

$$\begin{cases} \hat{\theta}(t) = \hat{\theta}(t-1) + K(t)\varepsilon(t), \\ \varepsilon(t) = y(t) - \hat{y}(t|t-k), \\ K(t) = P(t-1)\phi(t)[\lambda_t + \phi^T(t)P(t-1)\phi(t)]^{-1}, \\ P(t) = \lambda_t^{-1}[P(t-1) - K(t)(\lambda_t + \phi^T(t)P(t-1)\phi(t))^{-1}K^T(t)], \\ \phi(t) = (C(z^{-1})|_{\theta=\hat{\theta}(t)})^{-1}\varphi(t, \hat{\theta}(t)). \end{cases} \quad (28)$$

其中  $\lambda_t$  是遗忘因子, 实用时可取接近于 1 的常数, 或令  $\lambda_t = \alpha\lambda_{t-1} + (1 - \alpha)$ ,  $\alpha, \lambda_0 < 1$ <sup>[1]</sup>.

上述算法同样可以用于多步预报, 只需将参数集和数据集重新定义为

$$\begin{cases} \theta = (a_1, a_2, \dots, a_n; b_1, b_2, \dots, b_m; c_k, c_{k+1}, \dots, c_p)^T, \\ \varphi(t, \theta) = (-y(t+k-1|t, \theta), -y(t+k-2|t, \theta), \dots, \\ -y(t+k-n|t, \theta); u(t+k-1), u(t+k-2), \dots, \\ u(t+k-m); e(t), e(t-1), \dots, e(t+k-p))^T. \end{cases} \quad (29)$$

可以看出, 用直接法作多步预报, 对每一时刻都要用相同的格式反复作 1 步, 2 步, 直到  $k$  步的预报, 当  $k$  较大时计算量显然是很大的.

表 1 模拟试验两种预报形式结果的比较

$k = 1$ 步两种预报的比较				$k = 2$ 步两种预报的比较			
序号	输出实际值	单步预报	多步预报	序号	输出实际值	单步预报	多步预报
31	-2.21211	-2.16561	-2.16547	31	-4.28440	-4.71567	-4.71576
32	-3.53057	-3.53971	-3.54009	32	-5.56492	-5.65880	-5.65769
33	-5.38253	-5.25415	-5.25399	33	-6.89313	-6.99568	-6.99620
34	-7.25130	-7.18568	-7.18601	34	-5.86251	-5.83348	-5.83224
35	-6.55945	-6.67356	-6.67376	35	-2.31010	-2.70368	-2.70484
36	-3.12468	-3.19718	-3.19650	36	2.18768	2.04029	2.04195
37	1.41270	1.45568	1.45584	37	4.53944	4.46320	4.46326
38	3.74891	3.75040	3.75004	38	5.87533	5.58989	5.58947
39	5.11826	5.09432	5.09444	39	7.08459	7.18757	7.18908
40	6.61086	6.57233	6.57246	40	6.01399	6.15493	6.15406
41	5.92075	5.85852	5.85856	41	2.48238	2.50607	2.50573
42	2.76528	2.80052	2.80068	42	-0.03007	0.17280	0.17329
43	0.40894	0.52396	0.52377	43	-2.31057	-2.33562	-2.33666
44	-1.76899	-1.87645	-1.87689	44	-3.05616	-2.92010	-2.91940
45	-2.37687	-2.37651	-2.37577	45	-3.72146	-3.34497	-3.34546
46	-2.82080	-2.79129	-2.79151	46	-4.96787	-4.69673	-4.69674
47	-4.00901	-4.09971	-4.09965	47	-6.43760	-6.56720	-6.56819
48	-5.50166	-5.48846	-5.48813	48	-7.81214	-7.49256	-7.49048
49	-6.93709	-6.93673	-6.93689	49	-6.94793	-6.69144	-6.69397
50	-6.27572	-6.05233	-6.05231	50	-3.60364	-3.20464	-3.20380
预报误差方差		0.09203	0.09204	预报误差方差		0.21957	0.21956

## 五、结 论

ARMAX 模型的两种预报形式, 单步预报和多步预报, 不仅具有相同的预报误差最

小方差,而且是完全等价的。模拟试验表明二者的结果非常接近。实际应用时应根据问题的性质和需要决定用哪种形式。不论采用间接法或直接法进行计算,仅当预报步数  $k$  不大时多步预报的计算工作量较单步预报为小。

### 参 考 文 献

- [1] Ljung, L. and Söderström, T., Theory and Practice of Recursive Identification, The MIT Press, (1983).
- [2] Keyser, R. M. C. De and Cauwenberghe, A. R. Van, A Self-tuning Multistep Predictor Application, *Automatica*, 17(1981), 167—174.
- [3] Favier, G. and Dubois, D., A Review of  $k$ -step-ahead Predictors, *Automatica*, 26(1990), 75—84.
- [4] 邓自立等,多变量多步自校正递推预报器及其应用,自动化学报,9(1983),241—247.

## COMPARISON OF TWO FORMS OF ADAPTIVE $k$ -STEP-AHEAD PREDICTORS FOR ARMAX MODEL

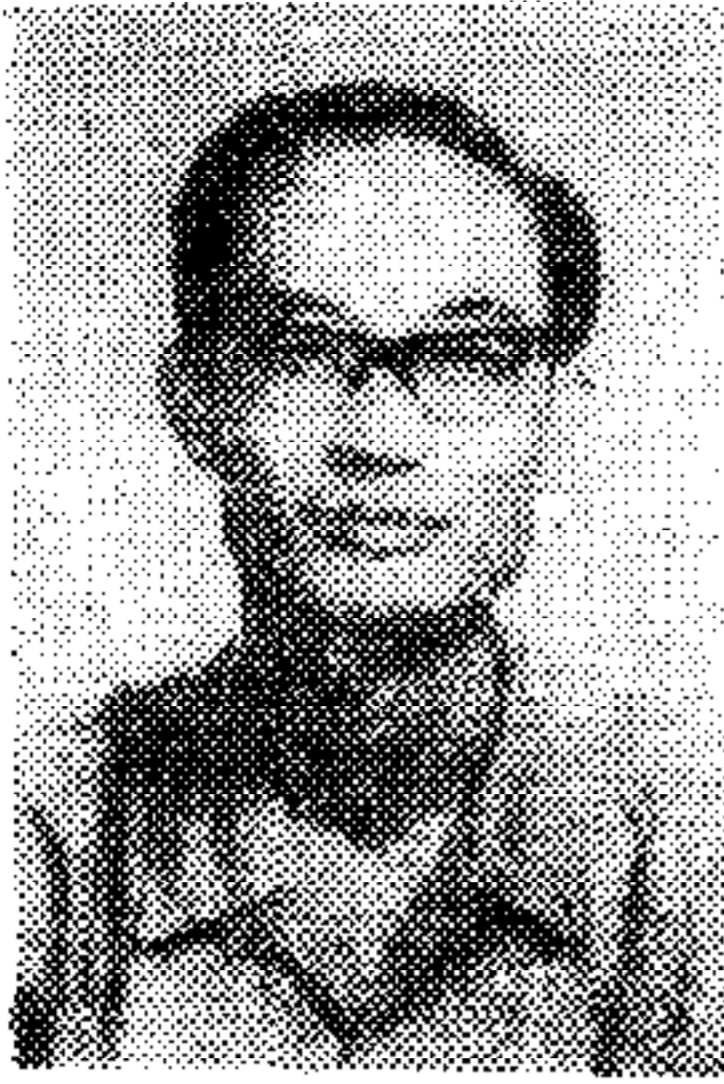
JIANG QIYUAN      SUN YIJUN

(Dept. of Applied Math., Tsinghua University, Beijing 100084)

### ABSTRACT

For  $k$ -step-ahead predictors of ARMAX model both single-step and multi-step forms are reviewed. The singlestep formulation depends on several past values of fixed step( $k$ -step) predictors, while the multistep form uses multiple recursivity of different step ( $k-1$ -step  $k-2$ -step, ...) predictors at fixed time. In this paper the equivalence relation of two forms is established, in a addition, the comparison of them from the viewpoint of application and simulation results is given.

**Key words:** ARMAX model;  $k$ -step-ahead predictor; singlestep form; multistep form.



**姜启源** 1939 年生于济南。1962 年毕业于清华大学。现为清华大学应用数学系运筹学与系统分析教研室主任,副教授,中国工业与应用数学学会数学模型专业委员会副主任。主要研究领域是数学建模,系统辨识及数学规划。



**孙义军** 1966 年 11 月生于大连。1991 年在清华大学应用数学系获硕士学位,从事系统辨识和计算机应用方面的研究工作。现在大连经济技术开发区三欧计算机有限公司任职。