We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

2

On-line Parameters Estimation with Application to Electrical Drives

Navid R. Abjadi[1], Javad Askari[1], Marzieh Kamali[1] and Jafar Soltani[2]
[1]Isfahan University of Tech., [2]Islamic Azad University- Khomeinishar Branch
Iran

1. Introduction

The main part of this chapter deals with introducing how to obtain models linear in parameters for real systems and then using observations from the system to estimate the parameters or to fit the models to the systems with a practical view.

Karl Friedrich Gauss formulated the principle of least squares at the end of the eighteenth century and used it to determine the orbits of planets and asteroids (Astrom & Wittenmark, 1995).

One of the main applications of on-line parameters estimation is self-tuning regulator in adaptive control; nevertheless other applications such as load monitoring or failure detection, estimation of some states to omit corresponding sensors and etc. also have great importance.

2. Models linear in parameters

A system is a collection of objects whose properties we want to study and a model of a system is a tool we use to answer questions about a system without having to do an experiment (Ljung & Glad, 1994). The models we work in this chapter are mathematical models, relationships between quantities.

There are different mathematical models categories such as (Ljung & Glad, 1994)

Deterministic-Stochastic

Stochastic models despite deterministic models contain stochastic variables or processes. Deterministic models are exact relationships between variables without uncertainty.

Dynamic-Static

The variables of a system usually change with time. If there is a direct, instantaneous relationship between these variables, the system or model is called static; otherwise the system is called dynamic. For example a resistor is a static system, but a series connection of a resistor and a capacitor is a dynamic system. In this chapter we interest dynamic systems which are described by differential or difference equations.

Continuous Time- Discrete Time

If the signals used in a model are continuous signals, the model is a continuous time model; which is described by differential equations. If the signals used in a model are sampled signals, the model is a discrete time model; which is described by difference equations.

Lumped-Distributed
Many physical systems are described by partial differential equations; the events in such systems are dispersed over the space variables. These systems are called distributed parameters systems. If a system is described by ordinary differential equations or a finite number of changing variables, it is a lumped system or model.

Change Oriented-Discrete Event Driven
The physical world and the laws of nature are usually described in continuous signals and variables, even discrete time systems obey the same basics. These systems are known as change oriented systems. For systems constructed by human, the changes take place in terms of discrete event, examples of such systems are queuing system and production system, which are called discrete event driven systems.

Models linear in parameters or linear regressions are among the most common models in statistics. The statistical theory of regression is concerned with the prediction of a variable y, on the basis of information provided by other measured variables φ_1, …, φ_n called the regression variables or regressors. The regressors can be functions of other measured variables. A model linear in parameters can be represented in the following form

$$y(t) = \varphi_1(t)\theta_1 + ... + \varphi_n(t)\theta_n = \varphi^T(t)\theta \tag{1}$$

where

$\varphi^T(t) = [\varphi_1(t) \ ... \ \varphi_n(t)]$, $\theta = [\theta_1 ... \theta_n]^T$ is the vector of parameters to be determined.

There are many systems whose models can be transformed to (1); including finite-impulse response (FIR) models, transfer function models, some nonlinear models and etc.

In some cases to attain (1), the time derivatives of some variables are needed. To avoid the noises in measurement data and to avoid the direct differentiation wich amplifies these noises, some filters may be applied on system dynamics.

Example: The d and q axis equivalent circuits of a rotor surface permanent magnet synchronous motor (SPMSM) drive are shown in Fig. 1. In these circuits the iron loss resistance is taken into account. From Fig. 1, the SPMSM mathematical model is obtained as (Abjadi et al., 2005)

$$\frac{di_{dm}}{dt} = -\frac{R}{K}i_{dm} + Pi_{qm}\omega_r + \frac{1}{K}v_d$$
$$\frac{di_{qm}}{dt} = -\frac{R}{K}i_{qm} - Pi_{dm}\omega_r - \frac{PK_\phi}{K}\omega_r + \frac{1}{K}v_q \tag{2}$$

where R, B, J, P and T_L are stator resistance, friction coefficient, momentum of inertia, number of pole pairs and load torque, also K and K_ϕ are defined by

$$K = (1+\frac{R}{R_i})L \quad , \quad K_\phi = (1+\frac{R}{R_i})\phi$$

here R_i, ϕ and L are respectively the motor iron loss resistance, rotor permanent magnet flux and stator inductance.
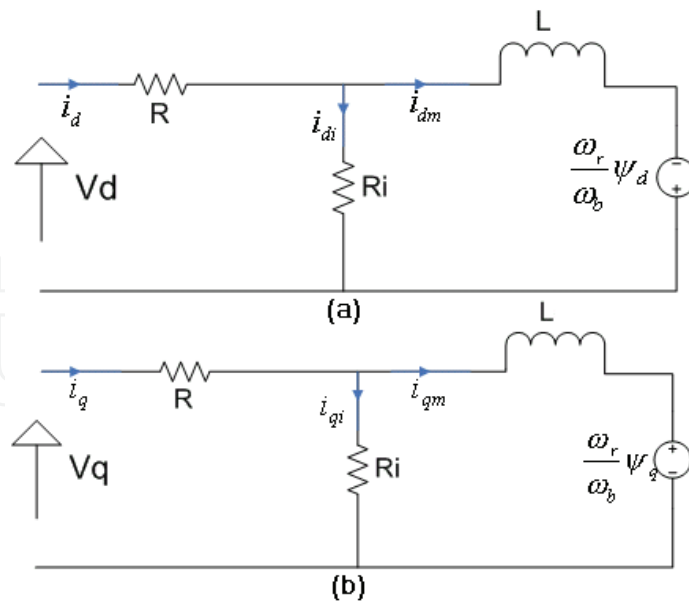
Figure 1. The d and q axis equivalent circuits of a SPMSM

From Fig. 1-b, the q axis voltage equation of SPMSM can be obtained as

$$Kpi_q = -Ri_q - KP\omega_r i_d - P\phi\omega_r + v_q + \frac{L}{R_i}pv_q + \frac{L}{R_i}P\omega_r v_d \tag{3}$$

where $p = \dfrac{d}{dt}$

Multiplying both sides of (3) by $\dfrac{1}{p+a}$, (3) becomes

$$K\frac{p}{p+a}i_q = -R\frac{1}{p+a}i_q - K\frac{1}{p+a}(P\omega_r i_d) - \phi\frac{1}{p+a}P\omega_r + \frac{1}{p+a}v_q$$
$$+ \frac{L}{R_i}\frac{p}{p+a}v_q + \frac{L}{R_i}\frac{1}{p+a}(P\omega_r v_d) \tag{4}$$

Assume

$$i_{qf} = \frac{1}{p+a}i_q, \ v_{df} = \frac{1}{p+a}v_d, \ \omega_{rf} = \frac{1}{p+a}\omega_r$$
$$\omega i_{df} = \frac{1}{p+a}\omega_r i_d, \ \omega v_{df} = \frac{1}{p+a}\omega_r v_d \tag{5}$$

then

$$\frac{p}{p+a}i_q = i_q - ai_{qf}, \ \frac{p}{p+a}v_q = v_q - av_{qf} \tag{6}$$

Linking (4), (5) and (6), yields

$$v_{qf} = K(i_q - ai_{qf} + P\omega i_{df}) + Ri_{qf} + \phi P\omega_{rf} - \frac{L}{R_i}(v_q - av_{qf} + P\omega v_{df}) \qquad (7)$$

Comparing (7) by (1), $y = v_{qf}$, $\theta = [K \ R \ \phi \ \frac{L}{R_i}]^T$,

$$\varphi^T = [i_q - ai_{qf} + P\omega i_{df} \quad i_{qf} \quad P\omega_{rf} \quad -(v_q - av_{qf} + P\omega v_{df})] .$$

3. Prediction Error Algorithms

In some parameters estimation algorithms, parameters are estimated such that the error between the observed data and the model output is minimized; these algorithms called prediction error algorithms. One of the prediction error algorithms is least squares estimation; which is an off-line algorithm. Changing this estimation algorithm to a recursive form, it can be used for on-line parameters estimation.

3.1 Least-Squares Estimation

In least square estimation, the unknown parameters are chosen in such a way that the sum of the squares of the differences between the actually observed and the computed (predicted) values, multiplied by some numbers, is a minimum (Astrom & Wittenmark, 1995).

Consider the models linear in parameters or linear regressions in (1), base on the least squares estimation the parameter θ are chosen to minimize the following loss function

$$J = \frac{1}{2} \sum_{t=1}^{N} w(t)[y(t) - \varphi^T(t)\hat{\theta}]^2 \qquad (8)$$

where $\hat{\theta}$ is the estimation of θ and $w(t)$ are positive weights.

There are several methods in literatures to obtain θ such that (8) becomes minimized, the first one is to expand (8), then separate it in two terms, one including θ (it can be shown this term is positive or equal to zero) the other independent of θ; by equating the first term to zero, (8) is minimized. In other approach the least squares problem is interpreted as a geometric problem. The observations vector is projected in the vector space spanned by regression vectors and then the parameters are obtained such that this projected vector is produced by a linear combination of regressors (Astrom & Wittenmark, 1995). The last approach which is used here to obtain estimated parameters is to determine the gradient of (8), since (8) is in a quadratic form by equating the gradient to zero, one can obtain an analytic solution as follow.

To simplify the solution assume

$$Y = [y(1) \ y(2) \dots y(N)]^T , \ E = [e(1) \ e(2) \dots e(N)]^T , \ \Phi = \begin{bmatrix} \varphi^T(1) \\ \vdots \\ \varphi^T(N) \end{bmatrix}^T$$

where $e(t) = y(t) - \varphi^T(t)\hat{\theta}$.

Using these notations on can obtain

$$E = Y - \Phi\hat{\theta} \tag{9}$$

then (8) can be rewritten as

$$J = \frac{1}{2}E^T W E \tag{10}$$

where W is a diagonal matrix of weights.
Substitute for E in (10)

$$J = \frac{1}{2}(Y - \Phi\hat{\theta})^T W (Y - \Phi\hat{\theta}) \tag{11}$$

Expand (11) and calculate its gradient with respect to $\hat{\theta}$

$$J = \frac{1}{2}Y^T W Y - Y^T W \Phi\hat{\theta} - \hat{\theta}^T \Phi^T W Y + \hat{\theta}^T \Phi^T W \Phi\hat{\theta} \tag{12}$$

$$\frac{\partial J}{\partial \hat{\theta}} = -Y^T W \Phi + \hat{\theta}^T \Phi^T W \Phi \tag{13}$$

Equating gradient to zero

$$\begin{aligned}
\hat{\theta} = \hat{\theta}(N) &= (\Phi^T W \Phi)^{-1} \Phi^T W Y \\
&= [\sum_{t=1}^{N} w(t)\varphi(t)\varphi^T(t)]^{-1} \sum_{t=1}^{N} w(t)\varphi(t)y(t)
\end{aligned} \tag{14}$$

provided that the inverse is existed; this condition is called an excitation condition.

Bias and Variance
There are two different source cause model inadequacy. One is the model error that arises because of the measurement noise and system noise. This causes model variations called variance errors. The other source is model deficiency, that means the model is not capable of describing the system. Such errors are called systematic errors or bias errors (Ljung & Glad, 1994).
The least-squares method can be interpreted in statistical terms. Assume the data are generated by

$$y(t) = \varphi^T(t)\theta + e(t) \tag{15}$$

where $\{e(t), \ t = 1, 2, ...\}$ is a sequence of independent, equally distributed random variables with zero mean. $e(t)$ is also assumed independent of $\varphi(t)$. The least-squares estimates are unbiased, that is, $E(\hat{\theta}(t)) = \theta$ and an estimate converges to the true parameter value as the number of observations increases toward infinity. This property is called consistency (Astrom & Wittenmark, 1995).

Recursive Least-Squares (RLS)

In adaptive controller such as self-tuning regulator the estimated parameters are needed on-line. The least-squares estimation in (14) is not suitable for real-time purposes. It is more convenient to convert (14) to a recursive form.

Define

$$P^{-1}(t) = \Phi^T(t)W(t)\Phi(t) = \sum_{i=1}^{t} w(i)\varphi\ (i)\varphi^T(i)$$
$$= P^{-1}(t-1) + w(t)\varphi\ (t)\varphi^T(t) \tag{16}$$

From (14)

$$\hat{\theta}(t-1) = P\ (t-1) \sum_{i=1}^{t-1} w(i)\varphi\ (i)y(i) \tag{17}$$

Expanding (14) and substituting for $\sum_{i=1}^{t-1} w(i)\varphi\ (i)y(i)$ from (17)

$$\hat{\theta}(t) = P(t)(\sum_{i=1}^{t-1} w(i)\varphi\ (i)y(i) + w(t)\varphi\ (t)y(t))$$
$$= P(t)(P^{-1}(t-1)\hat{\theta}(t-1) + w(t)\varphi\ (t)y(t)) \tag{18}$$

From (16) it follows that

$$\hat{\theta}(t) = P(t)((P^{-1}(t) - w(t)\varphi\ (t)\varphi^T(t))\hat{\theta}(t-1) + w(t)\varphi\ (t)y(t))$$
$$= \hat{\theta}(t-1) + P(t)w(t)\varphi\ (t)(y(t) - \varphi^T(t)\hat{\theta}(t-1)) \tag{19}$$

Using (16) and (19) together establish a recursive least-squares (RLS) algorithm. The major difficulty is the need of matrix inversion in (16) which can be solved by using matrix inversion lemma.

Matrix inversion lemma. Let A, C and $C^{-1} + DA^{-1}B$ be non-singular square matrices. Then

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1} \tag{20}$$

For the proof see (Ljung & Soderstrom, 1985) or (Astrom & Wittenmark, 1995). □

Applying this lemma to (16)

$$P\ (t) = [P^{-1}(t-1) + w(t)\varphi\ (t)\varphi^T(t)]^{-1}$$
$$= P\ (t-1) - P\ (t-1)\varphi\ (t)[\frac{1}{w(t)}I + \varphi^T(t)P\ (t-1)\varphi\ (t)]^{-1}\varphi^T(t)P\ (t-1) \tag{21}$$

Thus the formulas of RLS algorithm can be written as

$$\hat{\theta}(t) = \hat{\theta}(t-1) + P(t)w(t)\varphi\ (t)(y(t) - \varphi^T(t)\hat{\theta}(t-1))$$

$$P\ (t) = P\ (t-1) - P\ (t-1)\varphi\ (t)[\frac{1}{w(t)}I + \varphi^T(t)P\ (t-1)\varphi\ (t)]^{-1}\varphi^T(t)P\ (t-1)$$

(22)

It is worthwhile to note that if y is a scalar, $\frac{1}{w(t)}I + \varphi^T(t)P\ (t-1)\varphi\ (t)$ will be a scalar too

and there is no need to any matrix inversion in RLS algorithm.

In model (1), the vector of parameters is assumed to be constant, but in several cases parameters may vary. To overcome this problem, two methods have been suggested. First is to use a discount factor or a forgetting factor; by choosing the weights in (8) one can discount the effect of old data in parameters estimation. Second is to reset the matrix $P(t)$ alternatively with a diagonal matrix with large elements; this causes the parameters are estimated with larger steps in (22); for more details see (Astrom && Wittenmark, 1995).

Example: For a doubly-fed induction machine (DFIM) drive the following models linear in parameters can be obtained without and with considering iron loss resistance respectively (abjadi, et all, 2006)

Model 1.
$$y = v_{ds} - v_{dr}$$
$$\varphi^T = [i_{ds}, p i_{ds} - \omega_r i_{qs}, -i_{dr}, -p i_{dr}, -\omega_r i_{qs} - \omega_r i_{qr}]$$
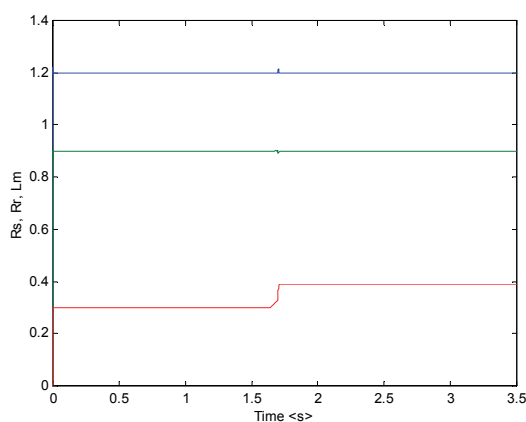$$\theta^T = [R_s, L_{ls}, R_r, L_{lr}, L_m]$$
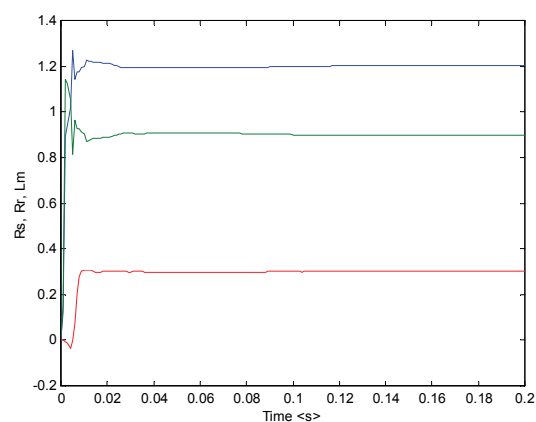
Model 2.
$$y = v_{ds} - v_{dr}$$
$$\varphi^T = [-p(v_{ds} - v_{dr}), p i_{ds}, i_{ds}, p^2 i_{ds} - p\omega_r i_{qs},$$
$$p i_{ds} - \omega_r i_{qs}, -\omega_r(i_{qs} + i_{qr}), -p i_{dr}, -i_{dr}, -p^2 i_{dr}]$$
$$\theta^T = [\frac{L_m}{R_i}, \frac{R_s L_m}{R_i}, R_s, \frac{L_{ls} L_m}{R_i}, L_{ls}, L_m, \frac{R_r L_m}{R_i} + L_{lr}, R_r, \frac{L_{lr} L_m}{R_i}]$$

(a) model 1   (b) model 2

Figure 2. Estimated parameters for DFIM

To solve the problem of derivatives ($p\,i_{ds}$, $p\,i_{dr}$) in model 1, a first order filter is used and in order to solve the problem caused by second derivatives in model 2, a second order filter is used.

The true parameters of the machine are given in Table 1. Using RLS algorithm, the estimated values of parameters are shown in Fig. 2. In Fig. 2.a. at the time t=1.65 s the value of the magnetizing inductance (L_m) increases 30 %. In this simulation the matrix $P(t)$ has been reset each 0.1 s with a diagonal matrix.

$P_n = 5.5\,kW$	$L_m = 300\,mH$
$R_s = 1.2\,\Omega$	$L_{ls} = 14\,mH$
$R_r = 0.9\,\Omega$	$L_{lr} = 12\,mH$

Table 1. Machine parameters

Simplified algorithms

There are simplified algorithms with less computation than RLS. Kaczmarz's projection algorithm is one of these algorithms. In this algorithm the following cost function is considered

$$J = \frac{1}{2}(\hat{\theta}(t) - \hat{\theta}(t-1))^T(\hat{\theta}(t) - \hat{\theta}(t-1)) + \alpha(y(t) - \varphi^T(t)\hat{\theta}(t)) \tag{23}$$

In fact in this algorithm $\hat{\theta}(t)$ is chosen such that $\left\| \hat{\theta}(t) - \hat{\theta}(t-1) \right\|$ is minimized subject to the constraint $y(t) = \varphi^T(t)\hat{\theta}(t)$. α is a Lagrangian multiplier in (23), taking derivatives with respect to $\hat{\theta}(t)$ and α the following parameters estimation law is obtained (Astrom & Wittenmark, 1995)

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\varphi(t)}{\varphi^T(t)\varphi(t)}(y(t) - \varphi^T(t)\hat{\theta}(t-1)) \tag{24}$$

To change the step length of the parameters adjustment and to avoid zero denominator in (24) the following modified estimation law is introduced

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\gamma\varphi(t)}{\lambda + \varphi^T(t)\varphi(t)}(y(t) - \varphi^T(t)\hat{\theta}(t-1)) \tag{25}$$

where $\lambda > 0$ and $0 < \gamma < 2$.

This algorithm is called normalized projection algorithm.

Iterative Search for Minimum

For many model structures the function $J = J(\hat{\theta})$ in (8) is a rather complicated function of $\hat{\theta}$, and the minimizing value must then be computed by computer numerical search for the minimum. The most common method to solve this problem is Newton-Raphson method (Ljung & Glad, 1994).

To minimize $J(\hat{\theta})$ its gradient should be equated to zero

$$\frac{\partial J(\hat{\theta})}{\partial \hat{\theta}} = 0 \tag{26}$$

It is achieved by the following recursive estimation

$$\hat{\theta}(t) = \hat{\theta}(t-1) - \mu(t-1)[J''(\hat{\theta}(t-1))]^{-1} J'(\hat{\theta}(t-1)) \tag{27}$$

Continuous-Time Estimation
Instead of considering the discrete framework to estimate parameters, one can consider continuous framework. Using analogue procedure similar parameter estimation laws can be obtained. For continuous gradient estimator and RLS see (Slotine & Weiping, 1991).

Model-Reference Estimation Techniques
Model-reference estimation techniques can be categorizes as techniques analog regression methods and techniques using Lyapunove or Passivity Theorem. For a detail discuss on techniques analog regression methods see (Ljung & Soderstrom, 1985) and for examples on Lyapunove or passivity theorem based techniques see (Soltani & Abjadi, 2002) & (Elbuluk, et all, 1998).

In model-reference techniques two models are considered; one contains the parameters to be determined (adaptive model) and the other is free or independent from those parameters (reference model). The two models have same kind output; a mechanism is used to estimate the parameters in such a way that the error between these models outputs becomes minimized or converges to zero.

3.2 Other Algorithms

Maximum Likelihood Estimation
In prior sections it was assumed that the observations are deterministic and reliable. But in stochastic studies, observations are supposed to be unreliable and are assumed as random variables. In this section we mention a method for estimating a parameter vector θ using random variables.

Consider the random variable $y = (y_1, y_2, ..., y_N) \in \Re^N$ as observations of the system. The probability that the realization indeed should take value y is described as $f(\theta; y)$, where $\theta \in \Re^d$ is the unknown parameter vector. A reasonable estimator for the vector θ is to determine it so that the function $f(\theta; y)$ takes it maximum (Ljung, 1999), i.e. the observed event becomes as likely as possible. So we can see that

$$\hat{\theta}_{ML}(y) = \underset{\theta}{\arg\max}\, f(\theta; y) \tag{28}$$

The function $f(\theta; y)$ is called the likelihood function and the maximizing vector $\hat{\theta}_{ML}(y)$ is known as the maximum likelihood. For a resistance maximum likelihood estimator and recursive maximum likelihood estimator see (Ljung & Soderstorm, 1985).

Instrumental Variable Method
Instrumental variable method, is a modification of the least squares method designed to overcome the convergence problems.
Consider the linear system

$$y(t) = \varphi^T(t)\theta + v(t) \tag{29}$$

In the least squares method, $\hat{\theta}(N)$ will not converge to θ, if there exists correlation between $\varphi(t)$ and $v(t)$ (Ljung, 1999). A solution for this problem is to replace $\varphi(t)$ by a vector $\zeta(t)$ that is uncorrelated with $v(t)$. The elements of $\zeta(t)$ are called instrumental variables and the estimation method is called instrumental variable method.

By replacing $\varphi(t)$ by $\zeta(t)$ in the least squares method we have

$$\hat{\theta}(N) = \left[\sum_{i=1}^{N} \zeta(t)\varphi^T(t) \right]^{-1} \sum_{i=1}^{N} \zeta(t)y(t) \tag{30}$$

for the off-line case and

$$\hat{\theta}(t) = \hat{\vartheta}(t-1) + L(t)[y(t) - \hat{\theta}^T(t-1)\varphi(t)],$$
$$L(t) = \frac{P(t-1)\zeta(t)}{1 + \varphi^T(t)P(t-1)\zeta(t)} = P(t)\zeta(t),$$
$$P(t) = P(t-1) - \frac{P(t-1)\zeta(t)\varphi^T(t)P(t-1)}{1 + \varphi^T(t)P(t-1)\zeta(t)}, \tag{31}$$

for recursive fashion.

The instrumental variables should be chosen such that

1. $\zeta(t)$ and $v(t)$ be uncorrelated,

2. The matrix $\lim\limits_{N \to \infty} \dfrac{1}{N} \sum\limits_{i=1}^{N} \zeta(t)\varphi^T(t)$ be invertible.

under these conditions and if $v(t)$ has zero mean, $\hat{\theta}(N)$ will converge to θ. A common choice of instrumental variables is (Ljung & Soderstorm, 1985)

$$\zeta^T(t) = (-y_M(t-1)...-y_M(t-n) \; u(t-1)...u(t-m)) \tag{32}$$

where $y_M(t)$ is the output of the system

$$y_M(t) + \bar{a}_1 y_M(t-1) + ... + \bar{a}_n y_M(t-n) = \bar{b}_1 u(t-1) + ... + \bar{b}_M u(t-m) \tag{33}$$

For the recursive fashion it is common to let $\bar{a}_i$ and $\bar{b}_i$ be time-dependent.

Bayesian Method

In the Bayesian method, in addition to observations, parameter is considered as a random variable too. In this method, parameter vector θ is considered to be a random vector with a certain prior distribution. The value of this parameter is determined using the observations u^t and y^t (input and output of the system until time t) of random variables that are correlated with it.

The posterior probability density function for θ is considered as $p(\theta | u^t, y^t)$. There are several ways to determine the parameter estimation $\hat{\theta}(t)$ from the posterior distribution. This is a very difficult problem in general to find the estimate $\hat{\theta}(t)$ and only approximate solutions can be found. But under the specific conditions mentioned in the following lemma, there exists an exact solution.

Lemma. (Ljung & Soderstorm, 1985) Suppose that the data is generated according to

$$y(t) = \varphi^T(t)\theta + e(t) \tag{34}$$

where the vector $\varphi(t)$ is a function of u^{t-1}, y^{t-1} and $\{e(t)\}$ is a sequence of independent Gaussian variable with $Ee(t) = 0$ and $Ee^2(t) = r_2(t)$. Suppose also that the prior distribution of θ is Gaussian with mean θ_0 and covariance matrix P_0. Then the posterior distribution $p(\theta | u^t, y^t)$ is also Gaussian with mean $\hat{\theta}(t)$ and covariance matrix $P(t)$, where $\hat{\theta}(t)$ and $P(t)$ are determined according to

$$
\begin{aligned}
&\hat{\theta}(t) = \hat{\theta}(t-1) + L(t)[y(t) - \hat{\theta}^T(t-1)\varphi(t)], \\
&L(t) = \frac{1}{r_2(t)} P(t)\varphi(t) = \frac{P(t-1)\varphi(t)}{r_2(t) + \varphi^T(t)P(t-1)\varphi(t)}, \\
&P(t) = P(t-1) - \frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{r_2(t) + \varphi^T(t)P(t-1)\varphi(t)}, \\
&\hat{\theta}(0) = \theta_0, \qquad P(0) = P_0
\end{aligned}
\tag{35}
$$

For the proof see (Ljung, 1985).

4. Nonlinear models

There are many applications that linear in parameters models dose not suffice to describe the system. Systems with nonlinearities are very common in real world; in this section some models suitable for such systems are introduced.

Wiener and Hammerstein System

Some especial cases of nonlinearities in system are static nonlinearities at the input or the output or both of them. In other words there are systems with dynamics with a linear nature, but there are static nonlinearities at the input or the output or both of them. Example for static nonlinearity at the input is saturation in the actuators and static nonlinearity at the output is sensors characteristics (Ljung, 1999).

A model with a static nonlinearity at the input is called a Hammerstein model while a model with a static nonlinearity at the output is called a Wiener model. Fig. 3 shows these models.
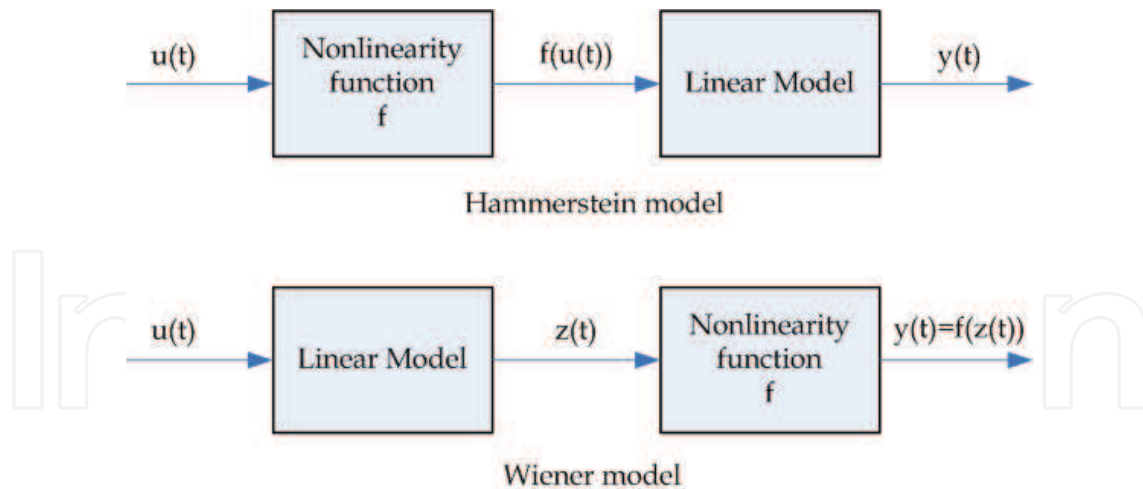
Figure 3. Hammerstein and Wiener models

Fuzzy System

Fuzzy Systems or knowledge-based systems are a type of nonlinear systems that can be used to approximate nonlinear behavior of many practical systems (Wang, 1997).

Certain types of fuzzy systems can be written as compact nonlinear formulas. In this section we will consider Takagi-Sugeno fuzzy systems that are a common used type of fuzzy systems.

Consider a multi input-single output Takagi-Sugeno fuzzy system given by (Passino & Yurkovich, 1998)

$$y = \frac{\sum\limits_{i=1}^{R} g_i(x)\mu_i(x)}{\sum\limits_{i=1}^{R} \mu_i(x)} \tag{36}$$

Where, $\mu_i(x)$ is the certainty of the premise of the i-th rule (Trabelsi & Lafont, 2004)

and $g_i(x) = a_{i,0} + a_{i,1}x_1 + ... + a_{i,n}x_n$ is the consequent of the i-th rule

With extending (36) we have

$$y = \frac{\sum\limits_{i=1}^{R} a_{i,0}\mu_i(x)}{\sum\limits_{i=1}^{R} \mu_i(x)} + \frac{\sum\limits_{i=1}^{R} a_{i,1}x_1\mu_i(x)}{\sum\limits_{i=1}^{R} \mu_i(x)} + ... + \frac{\sum\limits_{i=1}^{R} a_{i,n}x_n\mu_i(x)}{\sum\limits_{i=1}^{R} \mu_i(x)} \tag{37}$$

If we define

$$\xi_i(x) = \frac{\mu_i(x)}{\sum\limits_{i=1}^{R} \mu_i(x)} \tag{38}$$

$$\xi(x) = [\xi_1, \xi_2, ..., \xi_R, x_1\xi_1, x_1\xi_2, ..., x_1\xi_R, ..., x_n\xi_1, x_n\xi_2, ..., x_n\xi_R]^T \tag{39}$$

$$\theta = [a_{1,0}, a_{2,0}, \dots, a_{R,0}, a_{1,1}, a_{2,1}, \dots, a_{R,1}, \dots, a_{1,n}, a_{2,n}, \dots, a_{R,n}]^T \qquad (40)$$

We can write (36) as

$$y = \xi(x)^T \theta \qquad (41)$$

We see that (41) is in the same form as we defined for linear systems and is linear versus θ. Thus it is possible to use mentioned Estimators like recursive least square and Gradient Estimators to estimate parameter vector θ.

Neural Network System

Another common method to model systems is using artificial neural network for the details see (Ljung, 1999).

5. Conclusion

In this chapter, some parameters estimation algorithms are presented; among them RLS is one of the most common parameters estimation algorithms, which is discussed in details. The main key to use this algorithm or similar ones is the model must be linear in parameters.

Two practical examples from the field of electrical motor drives are introduced to show that even from nonlinear complex systems one may obtain models linear in parameters.

There are several ways to model a real system. Some of the models used to predict the behaviour of systems are presented.

6. References

Abjadi, N. R.; Soltani, J.; Pahlavaninezhad, M. & Askari, J. (2005). A Nonlinear Adaptive Controller for Speed Sensorless PMSM Taking The Iron Loss Resistance Into Account, *Proceedings of The Eighth International Conference on Electrical Machines and Systems (ICEMS)*, pp. 188-193, Nanjing, China, September 2005, Southeast University, Nanjing

Abjadi, N. R.; Askari, J. & Soltani, J. (2006). Adaptive Control of Doubly Fed Field-Oriented Induction Machine Based On Recursive Least Squares Method Taking the Iron Loss Into account, *Proceedings of CES/IEEE 5th International Power Electronics and Motion Control Conference (IPEMC)*, pp. 1923-1927, ISBN 1-4244-0448-7, Shanghai, China, August 2006, Shanghai Jiao Tong University, Shanghai

Astrom, K. J. & Wittenmark, B. (1995). *Adaptive control*, Addison-Wesley Longman, Inc., Californian

Elbuluk, M.; Langovsky, N. & Kankam, D. (1998). Design and Implementation of a Closed-Loop Observer and Adaptive Controller for Induction Motor Drives, *IEEE TRANSACTIONS ON INDUSTRY APPLICATIONS*, pp. 435-443, Vol. 34, No. 3, MAY/JUNE 1998

Ljung, L. & Soderstrom, T. (1985). *Theory and Practice of Recursive Identification*, MIT Press, Massachusetts

Ljung, L. & Glad, T. (1994). *Modeling of dynamic systems*, Prentice-Hall, Inc., ISBN 0-13-597097-0, NJ

Ljung, L. (1999). *System identification: Theory for the user,* Prentice-Hall, Inc., NJ

Passino, K. M. & Yurkovich S. (1998). *Fuzzy control,* Addison-Wesley Longman, Inc., Californian

Slotine, J. J. E. & Weiping, L. (1991). *Applied nonlinear control,* Prentice-Hall, Inc., NJ

Soltani, J. & Abjadi, N. R. (2002). A Modified Sliding Mode Speed Controller for an Induction Motor Drive without Speed Sensor Using the Feedback Linearization Theory, *Proceedings of EPE-PEMC 10th Power Electronics and Motion Control Conference,* Dubrovnik, Croatia, 2002, Zagreb University, Dubrovnik

Trabelsi, A.; Lafont F.; Kamoun M. & Enea G. (2004). Identification of nonlinear multivariable system by adaptive fuzzy Takagi-Sugeno model, *Int. J. Computational Cognition,* vol. 2, no. 3, pp. 137-153

Wang L. X. (1997). *A Course in Fuzzy Systems and Control,* Prentice-Hall, NJ

Frontiers in Adaptive Control

Edited by Shuang Cong

The objective of this book is to provide an up-to-date and state-of-the-art coverage of diverse aspects related to adaptive control theory, methodologies and applications. These include various robust techniques, performance enhancement techniques, techniques with less a-priori knowledge, nonlinear adaptive control techniques and intelligent adaptive techniques. There are several themes in this book which instance both the maturity and the novelty of the general adaptive control. Each chapter is introduced by a brief preamble providing the background and objectives of subject matter. The experiment results are presented in considerable detail in order to facilitate the comprehension of the theoretical development, as well as to increase sensitivity of applications in practical problems

INTECH
open science | open minds