

Combining Tracking and Regularization in Recursive Least Squares Identification

S.Gunnarsson ¹

Department of Electrical Engineering,
Linköping University,
S-58183 Linköping, Sweden
svante@isy.liu.se

Abstract

The combination of tracking and regularization in recursive identification is studied. It is shown that regularization of the information matrix corresponds to a normalization of the covariance matrix, and that several of the proposed methods for dealing with covariance matrix blow up can be interpreted as approximate implementations of covariance matrix normalization.

1 Introduction

In order for a recursive identification algorithm to have ability to track time varying systems and signals it is necessary to prevent the algorithm gain from tending to zero. This can be achieved by using, for example, exponential forgetting or covariance modification, see e.g. [1]. Such methods, however, make the algorithms sensitive for poor excitation. It therefore becomes necessary to introduce some safety mechanism that handles this situation. Several methods for dealing with this problem have been proposed. In [2], [3] and [4] different kinds of scalings of the covariance matrix are discussed. The results in [5] are closely related to the results that will be presented below, but in our approach the results are derived in a more straightforward way.

2 Recursive Parameter Estimation

We shall consider systems that can be described by a linear regression of the type

$$y_t = \varphi_t^T \theta_{t-1} + v_t \quad (1)$$

where y_t denotes the measured output signal and v_t is a disturbance. The regression vector φ_t contains delayed versions of the input signal u_t and the output signal y_t . Finally the vector θ_t contains the, possibly time varying, parameters of the system. For identification of the

parameters θ_t we shall consider algorithms of recursive least squares (RLS) type given by the structure

$$\hat{\theta}_t = \hat{\theta}_{t-1} + P_{t|t} \varphi_t (y_t - \varphi_t^T \hat{\theta}_{t-1}) \quad (2)$$

where $P_{t|t}$ is a symmetric matrix. We shall apply the terminology from state estimation, see [6], and split the update of $P_{t|t}$ into a measurement step and a time step. The measurement update is simplest formulated in terms of the information matrix $R_{t|t}$, defined by

$$R_{t|t} = P_{t|t}^{-1} \quad (3)$$

The measurement update is then given by, see [6],

$$R_{t|t} = R_{t|t-1} + \varphi_t \varphi_t^T \quad (4)$$

Applying the matrix inversion lemma gives the well known equation

$$P_{t|t} = P_{t|t-1} - \frac{P_{t|t-1} \varphi_t \varphi_t^T P_{t|t-1}}{1 + \varphi_t^T P_{t|t-1} \varphi_t} \quad (5)$$

Formulating the parameter estimation problem as the minimization of a weighted least squares criterion, see [1], the time update becomes

$$\bar{P}_{t+1|t} = \frac{1}{\lambda_t} P_{t|t} \quad (6)$$

where $0 < \lambda_t \leq 1$ is the forgetting factor, or equivalently

$$\bar{R}_{t+1|t} = \lambda_t R_{t|t} \quad (7)$$

By, on the other hand, assuming that the parameter vector of the true system varies according to a random walk the parameter estimation problem can be formulated as a state estimation problem, and the Kalman filter can be applied. The time update is then given by

$$\bar{P}_{t+1|t} = P_{t|t} + \Delta_t \quad (8)$$

where Δ_t is a symmetric matrix. Equation (8) together with the measurement update give, what is sometimes denoted, RLS with covariance resetting. Using the matrix inversion lemma equation (8) can be expressed, see [6], as

$$\bar{R}_{t+1|t} = R_{t|t} - R_{t|t} [R_{t|t} + \Delta_t^{-1}]^{-1} R_{t|t} \quad (9)$$

¹This work was sponsored by the Center for Industrial Information Technology (CENIIT) at Linköping University.

3 Regularization

A standard method for preventing the information matrix from becoming singular is to add a positive definite matrix to the information matrix to ensure that it always is invertible. This operation can be easily incorporated in a third update step given by

$$R_{t+1|t} = \bar{R}_{t+1|t} + \mu \cdot I \quad (10)$$

where μ is a positive scalar. Combining equations (9), (10) and (4) give

$$R_{t+1|t+1} = R_{t|t} - R_{t|t} [R_{t|t} + \Delta_t^{-1}]^{-1} R_{t|t} + \varphi_{t+1} \varphi_{t+1}^T + \mu \cdot I \quad (11)$$

in the Kalman filter case, while we for RLS with exponential forgetting obtain

$$R_{t+1|t+1} = \lambda_t R_{t|t} + \varphi_{t+1} \varphi_{t+1}^T + \mu \cdot I \quad (12)$$

This is the form a, so called, Levenberg-Marquardt regularization is carried out as discussed in, for example, [1]. It is obvious that we by adding the scaled identity matrix to the information matrix prevent it from becoming singular. The updating of the covariance matrix, corresponding to equation (10), is now obtained by applying the matrix inversion lemma. This yields

$$P_{t+1|t} = \bar{P}_{t+1|t} (I + \mu \bar{P}_{t+1|t})^{-1} \quad (13)$$

i.e. the regularization of the information matrix corresponds to a normalization of the covariance matrix. The covariance matrix update in the Kalman filter case is hence given by

$$\bar{P}_{t+1|t} = P_{t|t-1} - \frac{P_{t|t-1} \varphi_t \varphi_t^T P_{t|t-1}}{1 + \varphi_t^T P_{t|t-1} \varphi_t} + \Delta_t \quad (14)$$

together with the normalization in equation (13). Applying the same ideas to the RLS algorithm with exponential forgetting gives that the covariance matrix is given by

$$\bar{P}_{t+1|t} = \frac{1}{\lambda_t} (P_{t|t-1} - \frac{P_{t|t-1} \varphi_t \varphi_t^T P_{t|t-1}}{1 + \varphi_t^T P_{t|t-1} \varphi_t}) \quad (15)$$

together with the normalization in equation (13).

4 Related Algorithms

The normalization in equation (13) requires a matrix inversion and multiplication of two full rank matrices, and therefore a simpler operation could be of interest. One such simplification is achieved by replacing $\bar{P}_{t+1|t}$ in the second factor by the unit matrix. The normalization hence becomes

$$P_{t+1|t} = \frac{1}{(1 + \mu)} \bar{P}_{t+1|t} \quad (16)$$

i.e. the covariance matrix is scaled by a positive scalar less than one. Combining equation (16) with the time update in equation (8) yields

$$P_{t+1|t} = \frac{1}{(1 + \mu)} P_{t|t} + \frac{1}{(1 + \mu)} \Delta_t \quad (17)$$

Equation (17) together, the measurement update and the choice $\Delta_t = I$ is the, so called, Recursive Least Squares with Stabilized Forgetting, (RLS-SF) presented in [4]. This algorithm furthermore appears to be almost identical to the Selective Forgetting method (SF1) proposed in [3]. A related approach is to measure the magnitude of the matrix to be inverted in equation (13) by the trace of $\bar{P}_{t+1|t}$, and to replace equation (13) by

$$P_{t+1|t} = \frac{c}{\text{trace}(\bar{P}_{t+1|t})} \bar{P}_{t+1|t} \quad (18)$$

This then gives the constant-trace algorithm discussed in [2]. Finally, another way of approximating the normalization operation is to use that

$$(I + \mu \bar{P}_{t+1|t})^{-1} \approx I - \mu \bar{P}_{t+1|t} \quad (19)$$

for small μ . Inserted in equation (13) this gives

$$P_{t+1|t} = \bar{P}_{t+1|t} - \mu \bar{P}_{t+1|t}^2 \quad (20)$$

With some slight changes of time indices we get the Exponential Forgetting and Resetting Algorithm (EFRA), presented in [7], in which a term proportional to the square of the covariance matrix is subtracted from the covariance matrix.

References

- [1] L. Ljung and T. Söderström. *Theory and Practice of Recursive Identification*. M.I.T. Press, Cambridge, MA., 1983.
- [2] K. J. Åström and B. Wittenmark. *Adaptive Control*. Addison-Wesley, 1989.
- [3] Jens E. Parkum. *Recursive identification of time-varying systems*. PhD thesis, The Technical University of Denmark, Lyngby, Denmark, 1992.
- [4] J. J. Milek and F. J. Kraus. "Stabilized least squares estimators: Convergence and error propagation properties". In *Proc. 30th CDC*, pages 3086–3087, Brighton, England, 1991.
- [5] G. Kreisselmeier. "Stabilized least-squares type adaptive identifiers". *IEEE Trans. Automatic Control*, 35:306–310, 1990.
- [6] B. D. O. Anderson and J. B. Moore. *Optimal Filtering*. Prentice Hall, Englewood Cliffs, N.J., 1979.
- [7] M.E. Salgado, G.C. Goodwin, and R.H. Middleton. "Modified least squares algorithm incorporating exponential resetting and forgetting". *International Journal of Control*, 47:477–491, 1988.