Design of polynomial estimators based on covariances under multiple packet dropouts

Caballero-Águila, Raquel
Universidad de Jaén, Departamento de Estadística e I.O.
Campus Las Lagunillas, s/n
Jaén (23071),Spain
E-mail: raguila@ujaen.es

Hermoso-Carazo, Aurora and Linares-Pérez, Josefa
Universidad de Granada, Departamento de Estadística e I.O.
Campus Fuentenueva, s/n
Granada (18071),Spain
E-mail: ahermoso@ugr.es, jlinares@ugr.es

ABSTRACT

This paper is concerned with the least-squares polynomial filtering and fixed-point smoothing problems of discrete-time stochastic signals from observations featuring multiple packet dropouts, a realistic assumption in networked control systems and sensor networks where, generally, transmission losses are unavoidable due to the unreliable network characteristics. The technique used to derive the estimation algorithms consists of defining a suitable augmented observation model by assembling the original observations with their Kronecker powers, thus reducing the polynomial estimation problem from the original observations to the linear estimation problem from the augmented observations, which is addressed by an innovation approach. The proposed estimation algorithms do not require full knowledge of the state-space model generating the signal process, but only information about the dropout probabilities and the moments of the processes involved. To measure the performance of the estimators, recursive formulas for the filtering and fixed-point smoothing error covariance matrices are also proposed.

Introduction

Over the past decade, networked systems have attracted much attention due to their wide applicability in engineering problems. Nevertheless, such systems involve a number of inherent problems including (but not being limited to) communication time-delays and/or data packet dropouts, which are unavoidable due to numerous causes such as network congestion, random failures in the transmission mechanism, accidental loss of some measurements, or data inaccessibility at certain times.

These time-delays and packet dropouts are often random in nature, so they are modelled by introducing additional random variables in the observation model. Consequently, systems with random packet dropouts and/or time-delays are always non-Gaussian and, as in other kinds of non-Gaussian systems, the estimation problem has usually been focused on the search of suboptimal (basically linear) estimators. Under the assumption that the state-space model of the signal to be estimated is known, several modifications of conventional linear estimation algorithms have been proposed to incorporate the effects of random delays on the measurement arrival, and also many results have been reported on linear estimation for systems with packet dropouts (see e.g. [5] and references therein).

On another research line, there are many practical situations where the state-space model of the signal is not available and another type of information, for example about the covariance functions of the processes involved in the observation equation, must be processed for the estimation. In this context, linear estimation algorithms from randomly delayed observations based on covariance information have been derived under different hypotheses on the delay process and, also, polynomial estimators, which improve significantly the performance of linear ones, have been proposed in [1]. However, for systems with random packet dropouts, the estimation problem using covariance information have not been deeply studied yet.

This paper proposes a least-squares polynomial filtering and fixed-point smoothing algorithm from observations featuring multiple packet dropouts, which does not require knowledge of the state-space model but only information about the dropout probabilities and the moments of the processes involved. The measurement packet dropouts are modelled by introducing a sequence of Bernoulli random variables, whose values (one or zero) indicate if the current measure is available or lost during transmission (in whose case, the latest measurement is processed). The methodology used to address the polynomial estimation problem consists of augmenting the observation vectors by assembling the original vectors with their Kronecker powers. Then, by using an innovation approach, the polynomial estimators are derived as linear estimators from the augmented observations.

Observation model with multiple packet dropouts

Consider the problem of estimating an n-dimensional signal z_k, whose measured output $\widetilde{y}_k$ is perturbed by an additive noise v_k; that is,

(1) $$\widetilde{y}_k = z_k + v_k, \qquad k \geq 1.$$

Assume that, at the initial time $k = 1$, the measured output $\widetilde{y}_1$ is always available and, hence, the measurement processed for the estimation is equal to the real measurement, $y_1 = \widetilde{y}_1$. However, at any time $k > 1$, it is assumed that the measured output can be either transmitted successfully (with probability p_k) or dropped-out during transmission (with probability $1 - p_k$), in whose case the latest measurement received will be processed for the estimation.

This possibility of multiple random packet dropouts can be modelled by introducing a sequence of Bernoulli random variables, $\{\gamma_k;\ k > 1\}$, with $P[\gamma_k = 1] = p_k$, and considering the following model for the measurements processed to estimate the signal:

(2) $$y_k = \gamma_k \widetilde{y}_k + (1 - \gamma_k) y_{k-1}, \quad k > 1; \quad y_1 = \widetilde{y}_1.$$

Our aim is to obtain the least-squares (LS) νth-order polynomial estimator of the signal z_k based on the observations $y_1, \ldots, y_L$, with arbitrary integer order $\nu \geq 1$. Defining the random vectors

$$y_i^{[2]} = y_i \otimes y_i, \qquad y_i^{[j]} = y_i^{[j-1]} \otimes y_i, \quad j > 2$$

($\otimes$ denotes the Kronecker product (Magnus and Neudecker [3])), and assuming that $E\left[y_i^{[2\nu]}\right] < \infty$, this estimator is the orthogonal projection of z_k on the space of n-dimensional linear transformations of $y_1, \ldots, y_L$ and their Kronecker powers $y_1^{[2]}, \ldots, y_1^{[\nu]}, \ldots, y_L^{[2]}, \ldots y_L^{[\nu]}$. More specifically, we are interested in obtaining the LS νth-order polynomial filter ($L = k$) and fixed-point smoother ($L > k$) of the signal. For this purpose, we assume the following hypotheses on the signal and the noise processes involved in the observation model.

Model hypotheses

(H1) The $n \times 1$ signal process $\{z_k;\ k \geq 1\}$ has zero mean and, for $i, j = 1, \ldots, \nu$, the covariance function of the processes $\{z_k^{[i]};\ k \geq 1\}$ and $\{z_k^{[j]};\ k \geq 1\}$ can be expressed as

$$K_{k,s}^{ij} = E\left[\left\{z_k^{[i]} - E[z_k^{[i]}]\right\}\left\{z_s^{[j]} - E[z_s^{[j]}]\right\}^T\right] = A_k^{ij} B_s^{ijT}, \quad s \leq k$$

where A^{ij} and B^{ij} are $n^i \times N_{ij}$ and $n^j \times N_{ij}$ known matrix functions, respectively.

(H2) The noise process $\{v_k;\ k \geq 1\}$ is a zero-mean white sequence and its moments, up to the 2νth one, are known and will be denoted by

$$R_k^{ij} = E\left[\left\{v_k^{[i]} - E[v_k^{[i]}]\right\}\left\{v_k^{[j]} - E[v_k^{[j]}]\right\}^T\right],\ i, j = 1, \dots, \nu.$$

(H3) The noise $\{\gamma_k;\ k > 1\}$ is a sequence of independent Bernoulli random variables with known probabilities, $P[\gamma_k = 1] = p_k$.

(H4) The signal process, $\{z_k;\ k \geq 1\}$, and the noise processes, $\{v_k;\ k \geq 1\}$ and $\{\gamma_k;\ k > 1\}$, are mutually independent.

Polynomial estimation problem

Given the observation model (1)-(2), hypotheses (H1) and (H2) guarantee that $E\left[y_i^{[\nu]T} y_i^{[\nu]}\right] < \infty$ and, consequently, the LS νth-order polynomial estimator, $\hat{z}_{k/L}^{\nu}$, exists and can be obtained as the orthogonal projection of z_k on the space of n-dimensional linear transformations of $y_1, \dots, y_L$ and their Kronecker powers $y_1^{[2]}, \dots, y_1^{[\nu]}, \dots, y_L^{[2]}, \dots, y_L^{[\nu]}$. To obtain these estimators the following augmented observation vectors are defined by assembling the original vectors and their Kronecker powers,

$$\mathcal{Y}_k = \left(\ y_k^T, y_k^{[2]T}, \dots, y_k^{[\nu]T}\ \right)^T, \qquad \widetilde{\mathcal{Y}}_k = \left(\ \widetilde{y}_k^T, \widetilde{y}_k^{[2]T}, \dots, \widetilde{y}_k^{[\nu]T}\ \right)^T.$$

Clearly, the space of n-dimensional linear transformations of $\mathcal{Y}_1, \dots, \mathcal{Y}_L$ is equal to the space of n-dimensional linear transformations of $y_1, y_1^{[2]}, \dots, y_1^{[\nu]}, \dots, y_L, y_L^{[2]}, \dots, y_L^{[\nu]}$. Then, the LS νth-order polynomial estimator, $\hat{z}_{k/L}^{\nu}$, is the LS linear estimator of z_k based on $\mathcal{Y}_1, \dots, \mathcal{Y}_L$. To obtain this linear estimator, firstly the relation between the augmented vectors $\widetilde{\mathcal{Y}}_k$ and $\mathcal{Y}_k$ is studied, obtaining a new observation model with multiple packet dropouts, and secondly the second-order statistical properties of the processes involved in this new model are analyzed.

Augmented observation model

Using the binomial formula for the Kronecker powers (see Theorem 2.2.5 in [2]), the jth-order Kronecker powers of the real observations $\widetilde{y}_k$, given in (1), are written as

$$\widetilde{y}_k^{[j]} = \sum_{l=1}^{j} L_k^{jl} z_k^{[l]} + E[v_k^{[j]}] + g_k^j, \quad k \geq 1, \quad j = 1, \dots, \nu$$

where

$$L_k^{jl} = M_{j-l}^j(n)\left(E[v_k^{[j-l]}] \otimes I_{n,l}\right), \quad l < j; \qquad L_k^{jj} = I_{n,j}$$

and

$$g_k^j = \sum_{l=1}^{j-1} M_{j-l}^j(n)\Big(\left\{v_k^{[j-l]} - E[v_k^{[j-l]}]\right\} \otimes I_{n,l}\Big) z_k^{[l]} + \left\{v_k^{[j]} - E[v_k^{[j]}]\right\}$$

with

$$M_0^j(n) = M_j^j(n) = I_{n,j}, \quad M_r^j(n)(M_r^{j-1}(n) \otimes I_{n,1}) + (M_{r-1}^{j-1}(n) \otimes G_{j-r}), \quad r < j$$
$$G_l = (I_{n,1} \otimes G_{l-1})(G_1 \otimes I_{n,l-1}), \quad G_1 = K_{n,n}.$$

($I_{n,l}$ denotes the $n^l \times n^l$ identity matrix; $K_{n,n}$ is the $n^2 \times n^2$ commutation matrix, which satisfies $K_{n,n}(z_k \otimes v_k) = v_k \otimes z_k$).

Using again the Kronecker product properties and taking into account that $\gamma_k = \gamma_k^2 = \cdots = \gamma_k^\nu$, the following expression is derived for the jth-order Kronecker power of the available observations y_k, given by (2),

$$y_k^{[j]} = \gamma_k \widetilde{y}_k^{[j]} + (1-\gamma_k) y_{k-1}^{[j]}, \quad k>1; \quad y_1^{[j]} = \widetilde{y}_1^{[j]}, \quad j=1,\ldots,\nu.$$

Then, denoting

$$\mathcal{Z}_k = \begin{pmatrix} z_k \\ z_k^{[2]} \\ \vdots \\ z_k^{[\nu]} \end{pmatrix}, \quad \mathcal{L}_k = \begin{pmatrix} L_k^{11} & 0 & \cdots & 0 \\ L_k^{21} & L_k^{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ L_k^{\nu 1} & L_k^{\nu 2} & \cdots & L_k^{\nu\nu} \end{pmatrix}, \quad \mathcal{V}_k = \begin{pmatrix} 0 \\ E[v_k^{[2]}] \\ \vdots \\ E[v_k^{[\nu]}] \end{pmatrix}, \quad \mathcal{G}_k = \begin{pmatrix} v_k \\ g_k^2 \\ \vdots \\ g_k^\nu \end{pmatrix}$$

we obtain

$$\widetilde{\mathcal{Y}}_k = \mathcal{L}_k \mathcal{Z}_k + \mathcal{V}_k + \mathcal{G}_k, \qquad k \geq 1$$
$$\mathcal{Y}_k = \gamma_k \widetilde{\mathcal{Y}}_k + (1-\gamma_k)\mathcal{Y}_{k-1}, \quad k>1; \quad \mathcal{Y}_1 = \widetilde{\mathcal{Y}}_1.$$

Since the signal, $\mathcal{Z}_k$, and the noise, $\mathcal{V}_k + \mathcal{G}_k$, in this new model are non-zero mean vectors, we consider the centered vectors $\widetilde{Y}_k = \widetilde{\mathcal{Y}}_k - E\left[\widetilde{\mathcal{Y}}_k\right]$ and $Y_k = \mathcal{Y}_k - E[\mathcal{Y}_k]$, which, taking into account that $E\left[\mathcal{G}_k\right] = 0$, satisfy the following *augmented observation model*:

$$\widetilde{Y}_k = \mathcal{L}_k Z_k + \mathcal{G}_k, \qquad k \geq 1,$$
$$Y_k = \gamma_k \widetilde{Y}_k + (1-\gamma_k) Y_{k-1} + (\gamma_k - p_k) C_k, \quad k>1; \qquad Y_1 = \widetilde{Y}_1,$$

with $Z_k = \mathcal{Z}_k - E[\mathcal{Z}_k]$ and $C_k = E\left[\widetilde{\mathcal{Y}}_k\right] - E\left[\mathcal{Y}_{k-1}\right]$, where $E\left[\widetilde{\mathcal{Y}}_k\right]$ is calculated from

$$E\left[\widetilde{\mathcal{Y}}_k\right] = \mathcal{L}_k \begin{pmatrix} 0 \\ vec\left(A_k^{11} B_k^{11^T}\right) \\ \vdots \\ vec\left(A_k^{1\,\nu-1} B_k^{1\,\nu-1^T}\right) \end{pmatrix} + \begin{pmatrix} 0 \\ vec\left(R_k^{11}\right) \\ \vdots \\ vec\left(R_k^{1\,\nu-1}\right) \end{pmatrix}$$

and $E\left[\mathcal{Y}_k\right]$ is recursively obtained from

$$E\left[\mathcal{Y}_k\right] = (1-p_k) E\left[\mathcal{Y}_{k-1}\right] + p_k E[\widetilde{\mathcal{Y}}_k], \quad k>1; \quad E\left[\mathcal{Y}_1\right] = E[\widetilde{\mathcal{Y}}_1].$$

Note that the LS linear estimator of z_k based on $\mathcal{Y}_1, \ldots, \mathcal{Y}_L$ is equal to the LS linear estimator of z_k based on $Y_1, \ldots, Y_L$. To obtain this estimator, the second-order statistical properties of the augmented signal $\{Z_k;\ k \geq 1\}$ and augmented noise $\{\mathcal{G}_k;\ k \geq 1\}$ are necessary.

Statistical properties of the augmented processes

From the model hypotheses (H1)-(H4), the following statistical properties are derived (see Nakamori et al. [4]):

(P1) The augmented signal process $\{Z_k;\ k \geq 1\}$ has zero mean and its autocovariance function can be expressed in a semi-degenerate kernel form; namely

$$K_{k,s}^Z = E[Z_k Z_s^T] = \mathcal{A}_k \mathcal{B}_s^T, \quad s \leq k,$$

where

$$\mathcal{A}_k = \begin{pmatrix} A_k^{11} & \cdots & A_k^{1\nu} & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & A_k^{21} & \cdots & A_k^{2\nu} & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & A_k^{\nu 1} & \cdots & A_k^{\nu\nu} \end{pmatrix}$$

$$\mathcal{B}_k = \begin{pmatrix} B_k^{11} & \cdots & 0 & B_k^{21} & \cdots & 0 & \cdots & B_k^{\nu 1} & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \cdots & B_k^{1\nu} & 0 & \cdots & B_k^{2\nu} & \cdots & 0 & \cdots & B_k^{\nu\nu} \end{pmatrix}$$

(P2) The noise $\{\mathcal{G}_k;\ k \geq 1\}$ is a sequence of zero-mean mutually uncorrelated random vectors with covariance matrices $R_k^{\mathcal{G}} = E[\mathcal{G}_k \mathcal{G}_k^T]$, whose (r,s)-block is given by

$$\begin{aligned} R_k^{\mathcal{G}}(r,s) = &\ E\left[g_k^r g_k^{sT}\right] \\ = &\ \sum_{l=1}^{r-1}\sum_{i=0}^{s-1} M_{r-l}^r(n) P_{l,i}^{r,s}(v_k)(M_{s-i}^s(n))^T + \sum_{i=1}^{s-1} P_{0,i}^{r,s}(v_k)(M_{s-i}^s(n))^T + P_{0,0}^{r,s}(v_k) \end{aligned}$$

with

$$P_{l,i}^{r,s}(v_k) = vec^{-1}\left\{\left(I_{n,s-i} \otimes K_{n^i,n^{r-l}} \otimes I_{n,l}\right)\left(vec(R_k^{r-l\, s-i}) \otimes I_{n,l+i}\right) E[z_k^{[l+i]}]\right\}.$$

(P3) The noise process $\{\mathcal{G}_k;\ k \geq 1\}$ is uncorrelated with the signal process $\{Z_k;\ k \geq 1\}$, and both processes are independent of $\{\gamma_k;\ k > 1\}$.

Polynomial filtering and fixed-point smoothing algorithm

Using an innovation approach and the properties of the augmented processes (P1)-(P3) established above, we derive the following recursive algorithm, which provides the polynomial filtering and fixed-point smoothing estimators:

The polynomial fixed-point smoothers, $z_{k/L}^{\nu}$, are recursively obtained by

$$\hat{z}_{k/L}^{\nu} = \hat{z}_{k/L-1}^{\nu} + S_{k,L}\Pi_L^{-1}\mu_L, \qquad L > k$$

whose initial condition is given by the polynomial filter

$$\hat{z}_{k/k}^{\nu} = \mathcal{A}_k^{(1)} O_k, \quad k \geq 1,$$

where $\mathcal{A}_k^{(1)}$ is the submatrix constituted by the first n rows of $\mathcal{A}_k$.

The innovation, μ_k, satisfies

$$\mu_k = Y_k - p_k \mathcal{L}_k \mathcal{A}_k O_{k-1} - (1-p_k)Y_{k-1}, \quad k \geq 2; \quad \mu_1 = Y_1.$$

The vectors O_k are recursively calculated from

$$O_k = O_{k-1} + J_k \Pi_k^{-1}\nu_k, \quad k \geq 1; \qquad O_0 = 0.$$

The matrix function J_k satisfies

$$J_k = p_k\left[\mathcal{B}_k^T - r_{k-1}\mathcal{A}_k^T\right]\mathcal{L}_k^T, \quad k \geq 2; \quad J_1 = \mathcal{B}_1^T \mathcal{L}_1^T,$$

where r_k are recursively obtained from

$$r_k = r_{k-1} + J_k \Pi_k^{-1} J_k^T, \quad k \geq 1; \qquad r_0 = 0.$$

The covariance matrix of the innovation, Π_k, verifies

$$\begin{aligned} \Pi_k &= \Sigma_k^Y - p_k^2 \mathcal{L}_k \mathcal{A}_k r_{k-1} \mathcal{A}_k{}^T \mathcal{L}_k^T - (1-p_k)^2 \Sigma_{k-1}^Y \\ &\quad - p_k(1-p_k)\left[\mathcal{L}_k \mathcal{A}_k F_{k-1} + F_{k-1}^T \mathcal{A}_k{}^T \mathcal{L}_k^T\right], \quad k \geq 2; \\ \Pi_1 &= \Sigma_1^Y \end{aligned}$$

where Σ_k^Y *and* F_k *are recursively calculated from*

$$\Sigma_k^Y = p_k(\mathcal{L}_k\mathcal{A}_k\mathcal{B}_k^T\mathcal{L}_k^T + R_k^{\mathcal{G}}) + (1-p_k)\Sigma_{k-1}^Y + p_k(1-p_k)C_kC_k^T, \quad k \geq 2;$$
$$\Sigma_1^Y = \mathcal{L}_1\mathcal{A}_1\mathcal{B}_1^T\mathcal{L}_1^T + R_1^{\mathcal{G}}$$

and

$$F_k = J_k + p_k r_{k-1}\mathcal{A}_k^T\mathcal{L}_k^T + (1-p_k)F_{k-1}, \quad k \geq 2; \qquad F_1 = J_1.$$

Finally, the matrices $S_{k,L}$ *are calculated from*

$$S_{k,L} = p_L[\mathcal{B}_k^{(1)} - H_{k,L-1}]\mathcal{A}_L^T\mathcal{L}_L^T, \quad l > k; \qquad S_{k,k} = \mathcal{A}_k^{(1)}J_k,$$

where $\mathcal{B}_k^{(1)}$ *is the submatrix constituted by the first* n *rows of* $\mathcal{B}_k$*, and* $H_{k,L}$ *satisfy*

$$H_{k,L} = H_{k,L-1} + S_{k,L}\Pi_L^{-1}J_L^T, \quad L > k; \qquad H_{k,k} = \mathcal{A}_k^{(1)}r_k, \quad k \geq 1.$$

Estimation accuracy. The performance of the LS estimators $\hat{z}_{k/L}^{\nu}$, $L \geq k$ is measured by the covariance matrices of the estimation errors, $P_{k/L} = E\left[\left(z_k - \hat{z}_{k/L}^{\nu}\right)\left(z_k - \hat{z}_{k/L}^{\nu}\right)^T\right], \quad L \geq k$. These matrices are obtained by

$$P_{k/L} = P_{k/L-1} - S_{k,L}\Pi_L^{-1}S_{k,L}^T, \quad L > k$$

with initial condition

$$P_{k/k} = \mathcal{A}_k^{(1)}\left[\mathcal{B}_k^{(1)^T} - r_k\mathcal{A}_k^{(1)^T}\right].$$

Acknowledgments

This research is supported by Ministerio de Educación y Ciencia (grant No. MTM2008-05567) and Junta de Andalucía (grant No. P07-FQM-02701).

References

[1] Caballero-Águila R, Hermoso-Carazo A, Linares-Pérez J, Least-squares polynomial estimation from observations featuring correlated random delays, *Methodology and Computing in Applied Probability*, 12(3), 491–509 (2010).

[2] Carravetta F, Germani A, Raimondi M, Polynomial filtering for linear discrete time non-Gaussian systems, *SIAM J. Control and Optimization*, 34, 1666–1690 (1996).

[3] Magnus JR, Neudecker H Matrix Differential Calculus with Applications in Statistics and Econometrics (revised ed.). Wiley, New York (1999).

[4] Nakamori S, Caballero-Águila R, Hermoso-Carazo A, Jiménez López J, Linares-Pérez J, Polynomial fixed-point smoothing of uncertainly observed signals based on covariances, *International Journal of Systems Science*, 39(2), 207–216 (2008).

[5] Sun S, Xie L, Xiao W, Xiao N, Optimal filtering for systems with multiple packet dropouts, *IEEE Trans Circuits Syst II-Express Briefs*, 55(7), 695–699 (2008).