PAPER • OPEN ACCESS

Polynomial Wiener LQG Controllers based on Toeplitz Matrices

To cite this article: T. J. Moir 2022 *J. Phys.: Conf. Ser.* **2224** 012114

Polynomial Wiener LQG Controllers based on Toeplitz Matrices

T. J. Moir*

Department of Electrical and Electronic Engineering, School of Engineering, Computer and Mathematical Sciences (ECMS), Auckland, New Zealand.

*Email: tom.moir@aut.ac.nz

Abstract. This paper re-examines the discrete-time Linear Quadratic Gaussian (LQG) regulator problem. The normal approach to this problem is to use a Kalman filter state estimator and Kalman control state feedback. Though quite successful, an alternative approach in the frequency domain was employed later. That method used z-transfer functions or polynomials in the z-domain. The transfer function approach is similar to the method used in Wiener filtering and requires the use of Diophantine equations (sometimes bilateral) to find the optimal controller. The contribution here uses a similar approach but uses lower triangular Toeplitz matrices instead of polynomials to gain advantage of eliminating the use of Diophantine equations. This is because the single Diophantine equation approach fails when the system has non-relative prime polynomials and the need for bilateral Diophantine equations is computationally far more complex.

1. Introduction

The first approach to a generic state-space optimal control solution was obtained by Kalman[1] for the continuous time case. His method was based on minimising a quadratic cost function that penalises the error and output energy. It had a predecessor paper however in the form of a frequency-domain (also continuous-time) approach of Wiener [2], though Wiener only considered the optimal filtering case at the time. Whereas the Kalman approach needed a Ricatti equation to find the optimal state-feedback matrix, the Wiener approach found the transfer function of the optimal controller. The Wiener approach was not without its own difficulties in that it requires spectral factorization and the separation of causal from a mixture of causal and noncausal transfer functions to obtain a stable closed-loop system. Both approaches are essentially least-squares problems using differing system models. It was Kalman who first realised that the linear-quadratic (LQ) and the Linear-Quadratic Gaussian (LQG) problems could be found by treating the noise-free case and the filtering problems separately. Certainty equivalence as it was known means that for some time the LQG optimal control solution was that of a Kalman filter to estimates the noisy states followed by a state-feedback control law based on solving the LQ problem. The two problems were separable and the closed-loop poles were determined by a characteristic equation that involved both the Kalman filter and the state-feedback control-law[3, 4]. Kalman also introduced the idea of *controllability* and *observability.*

It took nearly another twenty years after Kalman's seminal work for the frequency-domain approach to regain popularity with the work of Shaked (continuous-time)[5, 6] and Grimble[7] for discrete-time. Later, Kucera developed polynomial methods involving Diophantine equations to ease the problem of obtaining the causal realisation of the controller[8]. The causal realisation is just a

Journal of Physics: Conference Series **2224** (2022) 012114 doi:10.1088/1742-6596/2224/1/012114

partial fraction expansion. Grimble introduced adaptive self-tuning methods using polynomial methods for the LQG case[9, 10]. Provided the polynomial of the plant are relative-prime (coprime) (a condition similar to controllability or observability in state-space approaches), only one 'implied' Diophantine equation[11] can be used to find a minimal degree solution. However, when the plant transfer function has common modes or there is an unstable signal source then two coupled Diophantine equations must be used[8, 12-14]. It is also worthy to note here that more advanced robust methods of controller design can also be solved using these Wiener polynomial methods using some novel extension methods[15].

We consider the scalar LQG problem as solved previously many times using polynomial methods. However, here we investigate the possibility of using lower-triangular Toeplitz (LTT) matrices instead of polynomials to represent a discrete-time system. The method uses a convolution matrix which is LTT and these behave in a similar manner to polynomials. Although the solutions that such an approach gives is finite-impulse in nature (i.e., an all-zero transfer-function), such methods often have advantages with stability over pole-zero based methods at the expense of using larger system models. The author has already applied this method to a new method of adaptive least-mean-squares (LMS)[16] and to Optimal filtering, predication and smoothing[17]. A similar approach has been used before[18, 19] for infinite dimensional Toeplitz spectra, that approach is more related to discrete state-space and not the polynomial Wiener LQG solution as given here.

2. Mathematical preliminaries

Before proceeding we need to consider discrete-time linear systems as applied to LTT matrices. We will show that the properties are almost identical to that of polynomials. We consider finite impulse response (FIR) transfer functions defined in negative powers of *z*, with polynomials of the form

$$w(z) = w_0 + w_1 z^{-1} + \dots + w_n z^{-n}, w_0 \neq 0 \tag{1}$$

considered a causal FIR transfer function with all (*n+1*) zeros of *z* lying within or on the unit circle of the z-plane. Likewise, the adjoint system

$$w\left(z^{-1}\right) = w_0 + w_1 z + \dots + w_n z^{n}, w_0 \neq 0 \tag{2}$$

has *(n+1)* zeros lying outside the unit circle in the z plane. The z-transform operator z^{-1} is commonly used interchangeably with the z-transform operator. A discrete time signal at time k can then be shifted one step backwards in time thus: $y_{k-1} = z^{-1} y_k$. The opposite is true for shifts forward in time by using z instead of its inverse.

Define the output y_k of a linear time-invariant system in terms of an input u_k

$$y_k = w(z) u_k \tag{3}$$

or in terms of the convolution summation

$$y_k = \sum_{i=0}^{k} w_{k-i} u_i \tag{4}$$

We can then write a matrix notation

$$\boldsymbol{y} = \boldsymbol{W}\boldsymbol{u} \tag{5}$$

where

$$\begin{bmatrix} y_0 \\ y_1 \\ . \\ . \\ y_m \end{bmatrix} = \begin{bmatrix} w_0 & 0 & . & . & 0 \\ w_1 & w_0 & 0 & . & 0 \\ w_2 & w_1 & w_0 & . & 0 \\ . & . & . & . & . \\ w_m & w_{m-1} & . & . & w_0 \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ . \\ . \\ u_m \end{bmatrix} \quad (6)$$

Bold letters here denote vectors or matrices. For convenience we omit the time dependence on k in the vectors. Note that $\boldsymbol{W}$ is a lower triangular square Toeplitz matrix (LTT) of dimension $m>n$. Such a matrix is characterized by the fact that each row is the previous one shifted to the right and a new value added in the preceding space of each row. The diagonal elements are all the same and all other elements are zero. Using this LTT matrix method a polynomial with a pure time-delay cannot be represented. Instead, a pure delay matrix must be cascaded with the LTT matrix. Consider a polynomial with a pure time-delay

$$\begin{aligned} y_k &= z^{-\ell} w(z) u_k \\ &= z^{-\ell}\left(w_0 + w_1 z^{-1} + \ldots + w_n z^{-n}\right) u_k \end{aligned} \quad (7)$$

We represent this in LTT form as follows

$$\boldsymbol{y} = \boldsymbol{W_D W u} \quad (8)$$

Suppose m=4, and the delay $\ell = 2$. Then $\boldsymbol{W}$ becomes generally

$$\boldsymbol{W} = \begin{bmatrix} w_0 & 0 & 0 & 0 & 0 \\ w_1 & w_0 & 0 & 0 & 0 \\ w_2 & w_1 & w_0 & 0 & 0 \\ w_3 & w_2 & w_1 & 0 & 0 \\ w_4 & w_3 & w_2 & w_1 & w_0 \end{bmatrix}$$

and we define the *delay* matrix as

$$\boldsymbol{W_D} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

If we multiply from the left or right, we obtain

$$\boldsymbol{W_D W} = \boldsymbol{W W_D} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ w_0 & 0 & 0 & 0 & 0 \\ w_1 & w_0 & 0 & 0 & 0 \\ w_2 & w_1 & w_0 & 0 & 0 \end{bmatrix}$$

Which has shifted the columns of $\boldsymbol{W}$ to the left by two columns and left two zero columns at the far right. We can do this for any order delay but clearly, we must choose $m >> \ell$ or too much information is lost. Also note that $\boldsymbol{W}_D$ is singular and cannot be inverted. The properties of LTT represented system are covered in more detail in references[17] and [20]. The essence is that when a

LTT is inverted, its contents become the power series expansion of the reciprocal original polynomial. For this to happen the roots of the polynomial must all lie within the unit circle of the z-plane. An unstable polynomial would give a divergent power-series expansion. This means that division of two polynomials become the multiplication of a LTT matrix with the inverse of the second. Furthermore, LTT matrices commute in the same way that multiplication of polynomials also commute. That is for two polynomials *a* and *b* we know that *ab=ba* and similarly their equivalent LTT matrices also commute in the form ***AB=BA***[21]**.** As we are dealing with Wiener type solutions, we also need a method for separating causal from noncausal systems. This usually occurs from a Laurent series. For example, for the Laurent series

$$g = g_{-3}z^3 + g_{-2}z^2 + g_{-1}z + g_0 + g_1 z^{-1} + g_2 z^{-2} + g_3 z^{-3}$$

A LTT matrix represents it below which can be split into causal and noncausal LTTs.

$$\boldsymbol{G} = \begin{bmatrix} g_0 & g_{-1} & g_{-2} & g_{-3} \\ g_1 & g_0 & g_{-1} & g_{-2} \\ g_2 & g_1 & g_0 & g_{-1} \\ g_3 & g_2 & g_1 & g_0 \end{bmatrix} = \boldsymbol{G}_+ + \boldsymbol{G}_-$$

Now a causal polynomial $g_+ = g_0 + g_1 z^{-1} + g_2 z^{-2} + g_3 z^{-3}$ and a noncausal polynomial $g_- = g_{-3}z^3 + g_{-2}z^2 + g_{-1}z$ are represented by their corresponding lower triangular Toeplitz matrices as follows:

$$\boldsymbol{G}_+ = \begin{bmatrix} g_0 & 0 & 0 & 0 \\ g_1 & g_0 & 0 & 0 \\ g_2 & g_1 & g_0 & 0 \\ g_3 & g_2 & g_1 & g_0 \end{bmatrix} \quad \boldsymbol{G}_- = \begin{bmatrix} 0 & g_{-1} & g_{-2} & g_{-3} \\ 0 & 0 & g_{-1} & g_{-2} \\ 0 & 0 & 0 & g_{-1} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We make sure that $\boldsymbol{G}_-$ has zero as the leading diagonal since the noncausal terms of a Laurent series have no term in z^0.

Spectral factorization follows in a similar manner. For example, a typical polynomial problem will produce a *symmetrical* factorization problem such as [22, 23]

$$w(z)qw(z^{-1}) + r = \Delta(z)\Delta(z^{-1}) \tag{9}$$

where *q* and *r* are noise variances (or control and output weightings) and *w(z)* is the signal generating system. The spectral factor $\Delta(z)$ is strictly Hurwitz so that its inverse is stable. Its adjoint spectral factor $\Delta(z^{-1})$ has all of its roots outside of the z-plane.

The polynomial multiplications $w(z)qw(z^{-1})+r$ give rise to a Laurent series which is *symmetrical* for negative and positive values of z.

In our Toeplitz matrix notation, we have an equivalent

$$\boldsymbol{WQW}^{\mathrm{T}} + \boldsymbol{R} = \boldsymbol{S} \tag{10}$$

where ***Q*** and ***R*** are noise covariance matrices and ***W*** is the signal generating matrix. The Matrix ***S*** is a full-rank Toeplitz Sylvester matrix[20, 24] formed from the *symmetric* Laurent series of $s = w(z)qw(z^{-1}) + r$ which is

$$s = s_{-m}z^m + \ldots + s_{-2}z^2 + s_{-1}z + s_0 + s_1 z^{-1} + s_2 z^{-2} + \ldots + s_m z^{-m} \tag{11}$$

$$\boldsymbol{S}=\begin{bmatrix} s_0 & s_1 & . & . & . & s_m & 0 & . & . & . & 0 \\ s_{-1} & s_0 & s_1 & . & . & . & & . & & & . \\ . & s_{-1} & . & . & & & & & . & & . \\ . & . & & . & . & & & & & . & . \\ . & . & & & . & . & & & & & 0 \\ s_{-m} & . & & & & & & & & & s_m \\ 0 & s_{-m} & & & & & & . & & & . \\ . & . & . & & & & . & . & . & & . \\ . & & . & . & & & & . & . & . & . \\ . & & & . & . & & & & . & . & s_1 \\ 0 & . & . & . & 0 & s_{-m} & . & . & . & s_{-1} & s_0 \end{bmatrix}. \tag{12}$$

Any spectral factorization must split **S** into a lower and upper-triangular Toeplitz matrix, since the result must represent a polynomial impulse-response. Fortunately such a technique has been known for some time as the method of Cholesky decomposition[23]. We factor accordingly.

$$\boldsymbol{S} = \boldsymbol{L}\boldsymbol{L}^{\mathrm{T}} \tag{13}$$

Here $\boldsymbol{L}$ is lower triangular Toeplitz. For such a factorization to exist, **S** must be full rank. We should also note that the diagonal elements of $\boldsymbol{L}$ are all unity so that it is unique. Ordinary, Cholesky factorization requires $\boldsymbol{O}\left(\frac{\mathrm{m}^3}{3}\right)$ operations, however fast methods using FFTs now exist to perform this factorization[24]. Clearly $\boldsymbol{L}$ represents the causal spectral factor, and its transpose the noncausal spectral factor. The spectral factor polynomial can be read off the bottom row of the matrix in reverse order as per our definition in (6). Although it is not by any means compulsory to use the Cholesky method, it seems obvious since the solution is given naturally in lower triangular Toeplitz format. Spectral factorization as applied to Toeplitz matrices was first explored by Pousson [25].

Autoregressive moving average (ARMA) models are also easily put into LTT form since the two transfer functions share the same denominator polynomial. For example, for the transfer function $y_k = \frac{d(z)}{a(z)}u_k$. Provided $a(z)$ is stable, it can be written $\boldsymbol{Ay} = \boldsymbol{Du}$. Then $\boldsymbol{y} = \boldsymbol{A}^{-1}\boldsymbol{Du} = \boldsymbol{W}_{\mathrm{n}}\boldsymbol{u}$. Hence the ARMAX model

$$y_k = z^{-\ell}\frac{b(z)}{a(z)}u_k + \frac{c(z)}{a(z)}\xi_k + v_k \tag{14}$$

which by using *spectral factorization* can be simplified to innovations format,

$$y_k = z^{-\ell}\frac{b(z)}{a(z)}u_k + \frac{d(z)}{a(z)}\varepsilon_k \tag{15}$$

can also be written in in the LTT form

$$\boldsymbol{y} = \boldsymbol{Wu} + \boldsymbol{W}_{\mathrm{n}}\boldsymbol{\varepsilon} \tag{16}$$

To be valid, the poles of $a(z)$ must all lie within the unit circle of the z plane and $\boldsymbol{W} = \boldsymbol{W}_{\mathrm{p}}\boldsymbol{W}_{\mathrm{D}}$ includes the time-delay.

3. LQG control problem

Figure 1 shows the regulator LQG problem in LTT format. All terms are LTT matrices and signals are vectors.

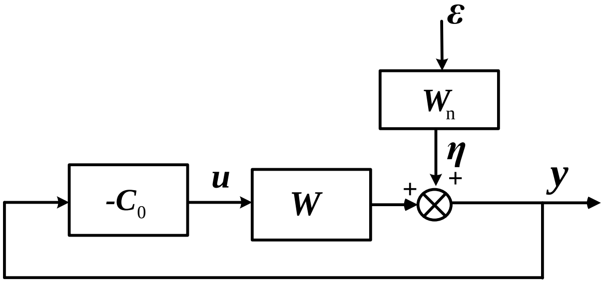

Figure 1. LQG Control problem

Minimise the LQG regulator criterion[10]

$$J = \boldsymbol{tr}E\left[q\boldsymbol{\Phi}_{yy} + r\boldsymbol{\Phi}_{\mathrm{uu}}\right] \tag{17}$$

The solution to the controller LTT matrix is found to be $\boldsymbol{C}_0 = \boldsymbol{M}\left[\boldsymbol{I} - \boldsymbol{W}_\mathrm{p}\boldsymbol{W}_\mathrm{D}\boldsymbol{M}\right]^{-1}$ where the closed loop optimal system is $\boldsymbol{M} = \boldsymbol{\Delta}^{-1}\left[\boldsymbol{\Delta}^{-\boldsymbol{T}}\boldsymbol{\Phi}_{\eta\eta}\boldsymbol{W}_\mathrm{p}^\mathrm{T}\boldsymbol{W}_\mathrm{D}^\mathrm{T}\boldsymbol{W}_\mathrm{n}^{-\mathrm{T}}\right]_+ \frac{q}{\sigma_\varepsilon^2}\boldsymbol{W}_\mathrm{n}^{-1}$ and the control spectral factor LTT satisfies $\boldsymbol{\Delta}^{\boldsymbol{T}}\boldsymbol{\Delta} = \boldsymbol{\Delta}\boldsymbol{\Delta}^{\boldsymbol{T}} = q\boldsymbol{W}^\mathrm{T}\boldsymbol{W} + r\boldsymbol{I}$.

Example: non-relative prime polynomials.

An example is now shown of an ARMAX model where $a(z) = \left(1\text{-}0.6z^{-1}\right)\left(1\text{-}0.7z^{-1}\right)$, $b(z) = z^{-2}\left(1\text{-}0.6z^{-1}\right)\left(1+2z^{-1}\right)$, $c(z) = \left(1\text{-}0.5z^{-1}\right)$ and all the process and measurement noise variances are unity. Consider the case when the weightings are *q=1, r=10.* The system has clearly non-relative prime polynomials and is non-minimum phase. As can be seen from Figure 2 such a difficulty does not pose any problem with this approach. Although mathematical models rarely suffer from this problem, when extended recursive least squares is used to estimate the parameters of an ARMAX model, if the system is over-parametrized, the problem of common factors (greatest common divisors) arises. Whilst the bilateral Diophantine approach can cope with this problem, the simpler and less computationally demanding one Diophantine solution (the so-called *implied* Diophantine equation) is singular and has no solution.

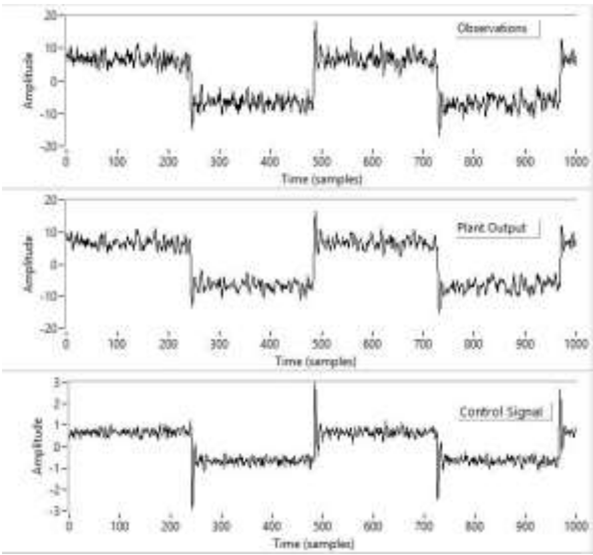

Figure 2. Non-relative prime polynomial example. Observations, (top) plant output (middle) and control signal (bottom) when *q=1, r=10.*

4.Conclusions

An alternative approach to LQG controller design has been shown using a lower triangular Toeplitz matrix method. The technique does not require Diophantine equations but relies on an FIR controller rather than the usual pole-zero approach. It will easily solve problems with systems that are non-relative prime but fails if the system is open-loop unstable.

Acknowledgement

I would like to thank Emeritus Professor M. J. Grimble of the University of Strathclyde, Scotland for helpful comments on Diophantine equations.

References

[1] R. E. Kalman 1960 Contributions to the theory of optimal control *Boletin de la Sociedad Mexicana,* **5** 102-119
[2] N. Wiener 1949 *Extrapolation, Interpolation and Smoothing of Stationary Time Series*. New York,Wiley
[3] H. Kwakernaak and R. Sivan 1972 *Linear optimal control systems*. New York: Wiley, 1972.
[4] B. D. O. Anderson and J. B. Moore 1971 *Linear optimal control*. Englewood Cliffs NJ: Prentice Hall
[5] U. Shaked 1976 A general transfer-function approach to the steady-state linear quadratic Gaussian stochastic control problem *International Journal of Control* **24**(6) 771-800
[6] J. F. Barrett 1975 Construction of linear-quadratic regulators using spectral factorization and the return-difference matrix *International Journal of Control* **26** 797-806
[7] M. Grimble 1979 Solution of the discrete-time stochastic optimal control problem in the z-domain *International Journal of Systems Science* **10**(12) 1369-1390
[8] V. Kučera 1993 Diophantine equations in control—a survey *Automatica* 29(6) 1361-1375
[9] M. J. Grimble and V. K. (eds) 1996 *Polynomial methods for control systems design*. NY: Springer-Verlag
[10] M. Grimble 1984 Implicit and explicit LQG self-tuning controllers *Automatica* **20**(5) 661-669
[11] V. Peterka 1972 On steady state minimum variance control strategy *Kybernetika* **8**(3) 219-232
[12] K. J. Hunt, M. ŠEbek, and M. J. Grimble 1987 Optimal multivariable LQG control using a single diophantine equation *International Journal of Control* **46**(4) 1445-1453
[13] M. J. Grimble 1986 Multivariate controllers for LQG self-tuning applications with coloured measurement noise and dynamic cost weighting *International Journal of Systems Science* **17**(4) 543-557
[14] M. J. Grimble 1988 Single versus double diophantine equation debate: comments on 'A polynomial approach to Wiener filtering' *International Journal of Control* 48(5) 2161-2165
[15] M. J. Grimble 1986 Optimal H∞ robustness and the relationship to LQG design problems *International Journal of Control,* **43**(2) 351-372
[16] T. Moir 2015 FIR system identification for correlated noise using the convolution matrix, an investigation *Signal, Image and Video Processing,* **2015** 1-6

[17] T. J. Moir 2018 Toeplitz matrices for LTI systems, an illustration of their application to Wiener filters and estimators *International Journal of Systems Science* 49(4) 800-817

[18] E. Jonckheere and L. H. Silverman 1978 Spectral theory of the linear-quadratic optimal control problem: Discrete-time single-input case *IEEE Transactions on Circuits and Systems* **25**(9) 810-825

[19] E.A. Jonckheere, J. Juang, and L.H. Silverman 1989 Spectral theory of the linear-quadratic and H infinity problems *Linear Algenra and its applications,* vol. 122-124, no. Sept-Nov, pp. 273-300

[20] U.Grenander and G.Szego 1958 *Toeplitz Forms and Their Applications*. University of California Press,Berkley

[21] D. Kucerovsky, K. Mousavand, and A. Sarraf 2016 On some properties of Toeplitz matrices *Cogent Mathematics* **3**(1) 1154705

[22] D. Youla and N. Kazanjian 1978 Bauer-type factorization of positive matrices and the theory of matrix polynomials orthogonal on the unit circle *IEEE Transactions on Circuits and Systems* **25**(2) 57-69

[23] V. Kucera 1979 *Discrete linear control, the polynomial equation approach.* John Wiley and Sons Ltd, New York

[24] M. Hromčík and M. Šebek 2007 Numerical algorithms for polynomial plus/minus factorization *International Journal of Robust and Nonlinear Control* **17**(8) 786-802

[25] H. R. Pousson 1968 Systems of Toeplitz Operators on H II *Transactions of the American Mathematical Society* **133**(2) 527-536