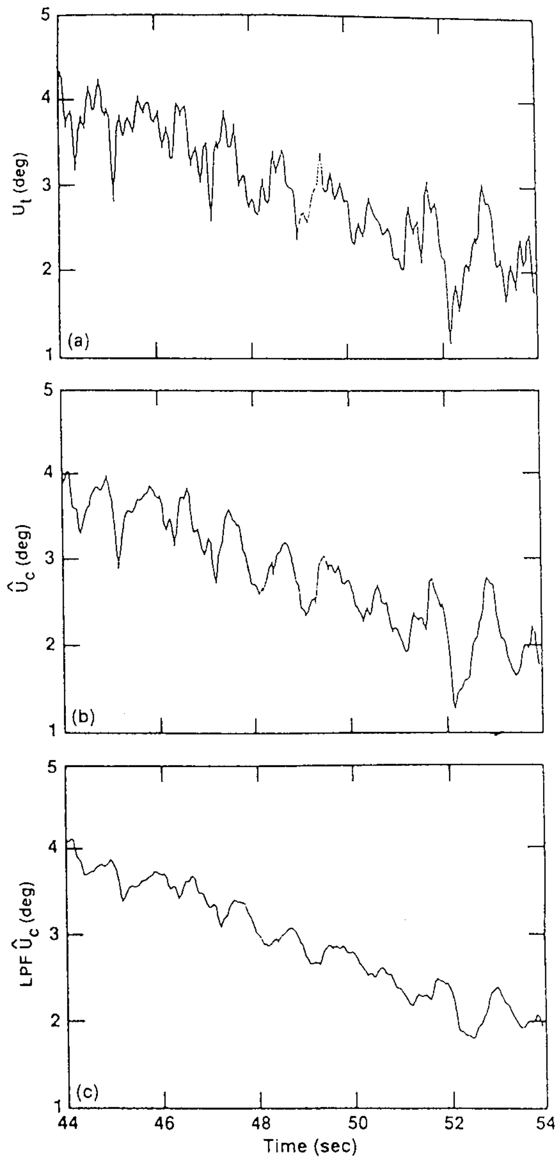

Fig. 2 Section of the time history during tracking of a continuously moving target: a) total head motion U_t; b) estimated voluntary head motion U_f; c) low-pass filter of U_f.

ing error. The use of the LPF-only configuration resulted in an increase of 33–56% in the dwelling time and in a decrease of 38–54% in the tracking error. Using the AF + LPF configuration increased the dwelling time by 30–60% and decreased the tracking error by 35–55%. Figure 2a shows the unfiltered signal, and Figs. 2b and 2c show the filtered signals. It can be seen that the AF filters the high-frequency periodic motion (Fig. 2b) and that the addition of the LPF attenuates the nonadditive biodynamic component (Fig. 2c).

From Table 2 one learns that the larger the random component in the vibration, the smaller is the contribution of the AF. The nonadditive component is the dominant biodynamic interference but cannot be handled by the adaptive filter since it is not correlated with the cabin vibration. In this regard, a distinction must be made between viewing tasks[7] and tracking tasks. In the viewing task the interference is additive and can be handled by noise cancellation methods. In the tracking task, however, the remnant noise increases with the intensity of the vibration and often becomes dominant. The remnant noise is not additive and cannot be directly reduced by the noise cancellation method.[7] Therefore, additional filtering schemes are needed to reduce the effects of biodynamic interference. For this reason, the subjects' performance with the LPF-only configuration was similar to their performance with the AF + LPF configuration. The experiments indicate that, at least with the time constants of 0.5 s, the subjects learned to compensate for the additional phase lag introduced by the LPF. As a rule, all of the subjects reached similar levels of performance, and eventually, after sufficient training, and with the adaptive and low-pass filtering configuration or with the low-pass-only configuration, it closely approached the tracking performance level without vibration.

Acknowledgments

This work was supported in part by the Human Factors Division, NASA Ames Research Center, and by the U.S. Air Force Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, under Grants NAWG 1128 and AFOSR88 0298, respectively.

References

[1]Levison, W. H., Baron, S., and Junker, A. M., "Modeling the Effects of Environmental Factors on Human Control and Information Processing," AMRL-76-74, Wright-Patterson AFB, OH, Aug. 1976.

[2]Jex, H. R., "Problems in Modeling Man Machine Control Behavior in Biodynamic Environments," *Proceedings of the 7th Annual Conference on Manual Control,* Scientific and Technical Information Office, NASA, Washington, DC, NASA SP-281, 1971, pp. 3–13.

[3]Wells, M. J., and Griffin, M. J., "Benefits of Helmet Mounted Display Image Stabilization Under Whole-Body Vibration," *Aviation, Space and Environmental Medicine,* Vol. 55, No. 1, Jan. 1984, pp. 13–18.

[4]Lewis, C. H., and Griffin, M. J., "Predicting the Effects of Vertical Vibration Frequency, Combination of Frequencies and Viewing Distance on the Reading of Numeric Displays," *Journal of Sound and Vibration,* 1980.

[5]Wells, M. J., and Griffin, M. J., "Performance with Helmet Mounted Sights," Univ. of Southampton, Southampton, England, UK, ISVR TR 152, Sept. 1987.

[6]Wells, M. J., and Griffin, M. J., "Tracking with the Head During Whole-Body Vibration," *Training, Human Decision Making and Control,* edited by J. Patrick and K. D. Duncan, Elsevier, New York, 1988.

[7]Lifshitz, S., and Merhav, S. J., "Adaptive Suppression of Biodynamic Interference in Helmet Mounted Displays and Head Teleoperation," *Journal of Guidance, Control, and Dynamics,* Vol. 14, No. 6, 1991, pp. 1173–1180.

[8]Tatham, N. O., "The Effect of Turbulence on Helmet Mounted Sight Aiming Accuracies," AGARD CP-267, March 1980.

Linear Quadratic Stochastic Control Using the Singular Value Decomposition

Yaakov Oshman*
Technion—Israel Institute of Technology, Haifa 32000, Israel

Introduction

THE conventional solution to the standard, discrete-time, linear quadratic Gaussian (LQG) stochastic control problem can be expressed in terms of the solution to two separate, dual problems: the linear quadratic optimal regulator problem and the linear optimal filtering problem. The inherent numerical instability of the discrete Riccati equation, which is solved in the Kalman filter via two covariance recursions (the time

Received May 9, 1990; revision received March 4, 1991; accepted for publication March 19, 1991.

*Senior Lecturer, Faculty of Aerospace Engineering. Member AIAA.

update and the measurement update), is now widely recognized. Since the dual optimal regulator problem involves the solution of a matrix Riccati equation that is identical in form to the filtering covariance recursion, the conventional method of computing the optimal control law suffers also from numerical instability as does the Kalman algorithm. This instability may cause severe problems, especially in ill-conditioned cases, e.g., where the model is poorly controllable.

Since square root algorithms have been useful in the past in overcoming the Kalman filter numerical instability,[1,2] this technique has also been applied to the dual control problem. Square root control algorithms that are based on UDU^T factorization[3] and Householder transformation[4] are available.

This Note presents a control law formulation that is based on the singular value decomposition (SVD). The new formulation is based on the observation that the Riccati recursion, which is the heart of the classical LQG solution, serves only as a means of computing the optimal gain matrix, and its solution is not needed in any other way. Thus, the optimal gain matrix is computed by the new algorithm without resorting to the solution of the Riccati equation. The new method complements the V-lambda square root filter,[2] which is also mainly based on the SVD, to form a complete, numerically robust and accurate LQG computational scheme.

The following linear, discrete-time, stochastic system is considered:

$$\mathbb{S}:\begin{cases} x_{k+1} = A_k x_k + B_k u_k + w_k \\ y_k = C_k x_k + v_k \end{cases} \tag{1}$$

for $k = 0, 1, \ldots, N-1$, where $x_k \in \Re^n$, $u_k \in \Re^p$, $y_k \in \Re^m$ and $\{w_k\}$, $\{v_k\}$ are the process and measurement zero mean Gaussian white sequences, respectively, and the initial state x_0 is a Gaussian random vector with mean m_0. The LQG stochastic control problem is to find the optimal control sequence $\{u_k^*\}_{k=0}^{N-1}$ which, based on the measurement history $\{y_k\}_{k=0}^{N-1}$, minimizes the cost functional

$$J = E\left\{\sum_{k=0}^{N-1}(x_{k+1}^T Q_k x_{k+1} + u_k^T R_k u_k)\right\} \tag{2}$$

Here $E(\cdot)$ denotes the expectation operator and $Q_k \geq 0$, and $R_k > 0$ are symmetric weighting matrices. As is well known, the optimal control strategy that minimizes the cost in Eq. (2) is a feedback control law that operates on an optimal estimate of the state, as follows:

$$u_k^* = -M_k \hat{x}_{k/k} \tag{3}$$

where $\hat{x}_{k/k}$ is the optimal state estimate (computed by the Kalman filter) and the gain matrix M_k is conventionally computed using the solution of the associated control Riccati recursion.

The new, SVD-based control law formulation is presented next.

New Control Law Formulation

Theorem: Given the dynamic system Eq. (1) and the cost functional Eq. (2), an SVD-based algorithm for the computation of the control gain in Eq. (3) is given by the following backward recursion for $k = N-1, N-2, \ldots, 0$:

Define the matrix $\Gamma_k \in \Re^{2n,n}$ as

$$\Gamma_k := \begin{bmatrix} Q_k^{T/2} \\ \Phi_{k+1} \end{bmatrix}, \qquad \Phi_N = 0 \tag{4}$$

where $(\cdot)^{1/2}$ denotes a lower triangular square root factor (e.g., a Cholesky factor) of $(\cdot)$ and $(\cdot)^{T/2} \equiv [(\cdot)^{1/2}]^T$. Perform a triangularization of Γ_k, i.e., find an orthogonal transformation Θ_k such that

$$\Theta_k \Gamma_k = \begin{bmatrix} \Pi_k \\ 0 \end{bmatrix} \tag{5}$$

where $\Pi_k \in \Re^{n,n}$ is upper triangular. Define the arrays $S_k \in \Re^{n+p,n}$, $T_k \in \Re^{n+p,p}$:

$$S_k := \begin{bmatrix} \Pi_k A_k \\ 0 \end{bmatrix}, \qquad T_k := \begin{bmatrix} \Pi_k B_k \\ R_k^{T/2} \end{bmatrix} \tag{6}$$

and perform an SVD of T_k to obtain

$$T_k = U_k \begin{bmatrix} \Sigma_k \\ 0 \end{bmatrix} V_k^T \tag{7}$$

Partition $U_k^T S_k$ in accordance with the partition of S_k in Eq. (6):

$$U_k^T S_k = \begin{bmatrix} \Psi_k \\ \Phi_k \end{bmatrix} \qquad \begin{matrix} \Psi_k \in \Re^{p,n} \\ \Phi_k \in \Re^{n,n} \end{matrix} \tag{8}$$

Then, the optimal control gain matrix at time k is given by

$$M_k = V_k \Sigma_k^{-1} \Psi_k \tag{9}$$

Moreover, defining J_k^*, the optimal cost-to-go, as

$$J_k^* := \min_{u_k, \ldots, u_{N-1}} \sum_{j=k}^{N-1} (\|x_{j+1}\|_{Q_j}^2 + \|u_j\|_{R_j}^2) \tag{10}$$

(where the symbol $\|z\|_A^2 := z^T A z$ is used), it is computed by

$$J_k^* = \|\Phi_k x_k\|^2 \qquad k = 0, 1, \ldots, N-1 \tag{11}$$

Proof: The required control law is also the optimal control law for the related deterministic system, where the random variables are replaced by their expected values, i.e., the following certainty equivalent system:

$$\mathbb{S}_{CE}:\begin{cases} x_{k+1} = A_k x_k + B_k u_k & x_0 = m_0 \\ y_k = C_k x_k \end{cases} \tag{12}$$

Hence, the following certainty equivalent cost functional will be minimized:

$$J_{CE} = \sum_{k=0}^{N-1}(x_{k+1}^T Q_k x_{k+1} + u_k^T R_k u_k) \tag{13}$$

Applying Bellman's dynamic programming principle of optimality, the theorem will be proved by induction.

Consider first the last stage of the process, assuming that the optimal control actions $\{u_0^*, u_1^*, \ldots, u_{N-2}^*\}$ have already been determined, so that u_{N-1}^* is the only control action yet to be found. By the optimality principle, the cost function to be minimized by u_{N-1}^* at the last stage is

$$J_{N-1} = \|x_N\|_{Q_{N-1}}^2 + \|u_{N-1}\|_{R_{N-1}}^2 \tag{14}$$

Using the certainty equivalent dynamic system Eq. (12) in Eq. (14) yields

$$J_{N-1} = \left\| \begin{bmatrix} Q_{N-1}^{T/2} A_{N-1} & Q_{N-1}^{T/2} B_{N-1} \\ 0 & R_{N-1}^{T/2} \end{bmatrix} \begin{bmatrix} x_{N-1} \\ u_{N-1} \end{bmatrix} \right\|^2 \tag{15}$$

Noting that $\Pi_{N-1} = Q_{N-1}^{T/2}$ and using the definitions in Eq. (6), Eq. (15) becomes

$$J_{N-1} = \|S_{N-1} x_{N-1} + T_{N-1} u_{N-1}\|^2 \tag{16}$$

Now perform a singular value decomposition of T_{N-1}:

$$T_{N-1} = U_{N-1} \begin{bmatrix} \Sigma_{N-1} \\ 0 \end{bmatrix} V_{N-1}^T \tag{17}$$

Using Eqs. (8) and (17) in Eq. (16) yields

$$J_{N-1} = \|\Sigma_{N-1} V_{N-1}^T u_{N-1} + \Psi_{N-1} x_{N-1}\|^2 + \|\Phi_{N-1} x_{N-1}\|^2 \tag{18}$$

Clearly, the minimum of J_{N-1} with respect to u_{N-1} is reached for

$$u^*_{N-1} = -V_{N-1}\Sigma^{-1}_{N-1}\Psi_{N-1}x_{N-1} \equiv -M_{N-1}x_{N-1} \tag{19}$$

and the minimization residual is

$$J^*_{N-1} = \min_{u_{N-1}} J_{N-1} = \|\Phi_{N-1}x_{N-1}\|^2 \tag{20}$$

Next, going backward to stage k of the process (where $0 \leq k < N-1$), it is assumed that all control actions prior to u_k have been determined, so that $\{u_k, \ldots, u_{N-1}\}$ are the only control actions yet to be exerted. By the principle of optimality, the optimal control action at stage k is determined by minimizing the following cost function (the cost-to-go) with respect to u_k, subject to the constraint in Eq. (12):

$$J_k = \|x_{k+1}\|_{Q_k} + \|u_k\|_{R_k} + J_{k+1}(x_{k+1}) \tag{21}$$

Noting that, by the induction assumption,

$$J^*_{k+1} = \|\Phi_{k+1}x_{k+1}\|^2 \tag{22}$$

and employing the definition in Eq. (4) of Γ_k, Eq. (21) can be rewritten as

$$J_k = \|\Gamma_k x_{k+1}\|^2 + \|R_k^{T/2}u_k\|^2 \tag{23}$$

Now find an orthogonal matrix Θ_k such that $\Theta_k\Gamma_k$ is upper triangular, i.e.,

$$\Theta_k\Gamma_k = \begin{bmatrix}\Pi_k \\ 0\end{bmatrix}, \qquad \Pi_k \text{ upper triangular} \tag{24}$$

(Θ_k need not be computed explicitly). Employing Eq. (24) and using the system Eq. (12) to express x_{k+1} as a function of x_k and u_k in Eq. (23) yields

$$J_k = \left\| \begin{bmatrix}\Pi_k A_k & \Pi_k B_k \\ 0 & R_k^{T/2}\end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix} \right\|^2 \tag{25}$$

which, using the definition in Eq. (6) of S_k, T_k, can be expressed as

$$J_k = \|S_k x_k + T_k u_k\|^2 \tag{26}$$

The last equation has the same form as that of Eq. (16). Replacing the index $N-1$ by k and following along the lines of derivation in the first part of the proof finally yields

$$u^*_k = -V_k\Sigma_k^{-1}\Psi_k x_k \equiv -M_k x_k \tag{27}$$

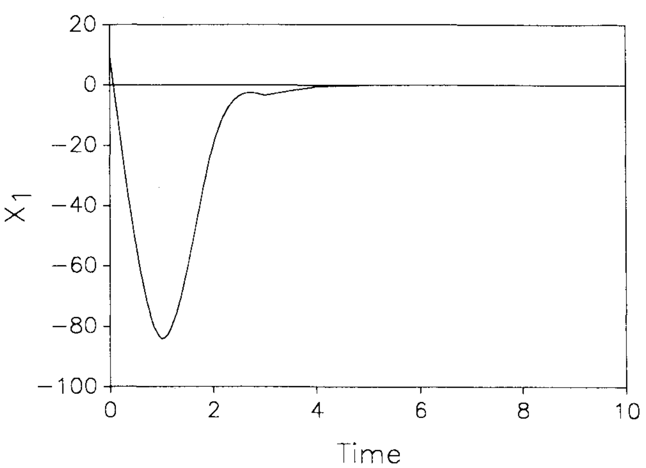

Fig. 1 Closed-loop response of the first state component (new algorithm in single precision, conventional algorithm in double precision).

The minimization residual (the optimal cost-to-go) is

$$J_k = \min_{u_k, \ldots, u_{N-1}} J_k = \|\Phi_k x_k\|^2 \tag{28}$$

which completes the proof. ■

Numerical Example

The following certainty equivalent dynamic system is considered:

$$x_{k+1} = \begin{bmatrix} 8.25 & 0.0 & 0.1 \\ 0.1 & 1.495 & 0.5 \\ 0.0 & 0.0 & 8.75 \end{bmatrix} x_k + \begin{bmatrix} 0.0 & 1.0 \\ 1.0 & 0.0 \\ 0.0 & 1.0 \end{bmatrix} u_k \tag{29}$$

$$x_0 = [10.0 \quad 10.0 \quad 10.0]^T$$

The cost functional is

$$J = \sum_{k=0}^{199} (x^T_{k+1}\,\mathrm{diag}\{0.5,\, 0.1,\, 0.5\}x_{k+1} + u_k^T\,\mathrm{diag}\{10.0,\, 10.0\}u_k)$$

The plant, Eq. (29), is unstable, but controllable. An Intel 80386/387 CPU/FPU-based Olivetti M380/C computer was used for the simulation. All programs were written in Microsoft Fortran 4.1, using LINPACK[5] mathematical routines. The conventional LQG algorithm was run in both single precision (SP) and double precision (DP). However, only the DP run was successful, while the SP solution diverged before completing the backward recursion. When the new algorithm was used in SP, no such difficulties were observed. The optimal gain sequence computed was identical to that obtained via the DP conventional method. In Fig. 1, the time history of the closed-loop system's first state component is shown. As can be observed, the system's response during the initial time steps is violent, which might explain the divergence of the SP version of the Riccati recursion. It is emphasized, however, that no such problems were encountered when the new, SVD-based algorithm was used.

Concluding Remarks

Inheriting the excellent numerical characteristics of the SVD, the new algorithm is guaranteed to be numerically stable and highly accurate. Moreover, cases of singular weighting matrices (i.e., unconstrained control action or unweighted states) can be handled without any modification. Combining the new control algorithm with the V-lambda square root filter, which also relies mainly on the SVD procedure, renders the resulting LQG scheme numerically robust and simple to implement in practice.

Acknowledgment

This work was supported by the Technion V.P.R. Fund-Sophia L. Shamban Research Fund.

References

[1]Bierman, G. J., *Factorization Methods for Discrete Sequential Estimation,* Academic Press, New York, 1977.

[2]Oshman, Y., "Gain-Free Square Root Information Filtering Using the Spectral Decomposition," *Journal of Guidance, Control, and Dynamics,* Vol. 12, No. 5, 1989, pp. 681-690.

[3]Thornton, C. L., and Jacobson, R. A., "Linear Stochastic Control Using the UDU^T Matrix Factorization," *Journal of Guidance and Control,* Vol. 1, No. 4, 1978, pp. 232-236.

[4]Silverman, L. M., "Discrete Riccati Equations: Alternative Algorithms, Asymptotic Properties, and System Theory Interpretations," *Control and Dynamic Systems—Advances in Theory and Applications,* edited by C. T. Leondes, Vol. 12, Academic Press, New York, 1976, pp. 313-386.

[5]Dongarra, J. J., Moler, C. B., Bunch, J. R., and Stewart, G. W., *LINPACK—User's Guide,* Society for Industrial and Applied Mathematics, Philadelphia, PA, 1979.