Some Formal Effects of Nonlinearities in Optimal Perturbation Control

Warren W. Willman*
U.S. Naval Research Laboratory, Washington, D.C.

Abstract

ONE stochastic optimal control problem of considerable interest is that of minimizing the expected value of a quadratic performance criterion in the presence of linear dynamics and state measurements which are both perturbed by Gaussian white noise. This case is important partly because it is the form of the first-order description of noise-induced perturbations from nominal behavior in a wider class of optimal control situations.[1] If the description of these perturbations is carried out to one higher order of accuracy, the effect is typically to introduce quadratic terms in the dynamics and state measurements and cubic terms in the performance criterion. The resulting control problem can often be rescaled so that the state, control, and measurement perturbations are of order unity and the *coefficients* of the added higher-degree terms become the relatively small quantities. An approximation to the optimal control law which is formally accurate to first order in these small coefficients is presented here for a class of problems with this latter type of structure. For such problems with the preceding origin, this degree of accuracy is all that is consistent with that of the problem formulation. Approximate solutions to similar modifications of the standard "linear-quadratic-Gaussian" control problem have been derived elsewhere under various assumptions.[2,3] The basic assumption made here is that the neglected error terms in the approximations are sufficiently "well behaved" that all quantities that *appear* to be of higher order *are* so in some appropriate sense; i.e., the results are only formal.

Contents

Unless otherwise indicated, lower case letters denote (real) column vectors or scalars. Matrices are denoted by capital Roman letters. A^T denotes the transpose of a matrix A. If A is square, $\mathrm{Tr}(A)$ denotes its trace and $|A|$ its determinant. Three-way matrices are denoted by capital Greek letters, and the following definitions are adopted for such an object Γ, with vector x and matrices A and B of compatible dimensions, and with repeated indices denoting summation:

$(\Gamma x)_{ij}$	$=\Gamma_{ij\sigma}x_\sigma$	(matrix)
$(Ax^T)_{ijk}$	$=A_{ij}x_k$	(three-way matrix)
$(A\Gamma)_{ijk}$	$=A_{i\sigma}\Gamma_{\sigma jk}$	(three-way matrix)
$(\Gamma B)_{ijk}$	$=\Gamma_{ij\sigma}B_{\sigma k}$	(three-way matrix)
$(\Gamma')_{ijk}$	$=\Gamma_{jki}$ and $(\Gamma^T)_{ijk}=\Gamma_{kji}$	(three-way matrices)
$[\mathrm{Tr}(\Gamma)]_i$	$=\Gamma_{\sigma i\sigma}$	(column vector)
$[\mathrm{Tr}(\Gamma\Gamma)]_{ij}$	$=\Gamma_{\sigma i\lambda}\Gamma_{\lambda j\sigma}$	(matrix)

Γ is called symmetric if $\Gamma=\Gamma'=\Gamma''=\Gamma^T$

With these definitions, the expression $AFTBDxx^T$ is fully associative.

We consider a stochastic optimal control problem in which, except perhaps for extremely large variable values, the dynamics are (in the sense of Ito equations[4])

$$dx=[Fx+Gu+\mathrm{Tr}(\Delta xx^T+2\Theta''xu^T)]dt + (I+\bar{\Psi}'x+\bar{\Gamma}'u)^T dw \tag{1}$$

where $x(0)$ is Gaussian with mean $\hat{x}_0$ and covariance P_0, and the (scalar) performance criterion is

$$J=E\Big\{\tfrac{1}{2}x_f^TS_fx_f+\tfrac{1}{3}x_f^T\mathrm{Tr}(\Pi_fx_fx_f^T)+\int_0^{t_f}[\tfrac{1}{2}(x^TAx+u^TBu) + c^Tu+\tfrac{1}{3}x^T\mathrm{Tr}(\Sigma xx^T)+\tfrac{1}{3}u^T\mathrm{Tr}(\Xi uu^T)]dt\Big\} \tag{2}$$

where t_f is fixed and E denotes prior expected value. The available state measurements are discrete samples $y(i\epsilon)$, $i=1,2,\ldots$, of a continuous-time process y for which

$$dy=[Hx+\mathrm{Tr}(\Phi xx^T)]dt+(I+\bar{\Upsilon}'x+\bar{\Omega}'u)^T dv \tag{3}$$

where x is the state vector, u is the control vector, and w and v are independent Wiener processes with covariance parameters $Q(t)$ and $R(t)$. The time argument t of these variables and of the coefficient matrices, which may be time-varying, is suppressed in the notation. The matrices A, Q, and S_f are symmetric and positive-semidefinite; B and R are symmetric and positive-definite. The three-way matrices Π_f, Σ and Ξ are symmetric, $\Delta=\Delta^T$ and $\Phi=\Phi^T$. The components of c and the three-way matrices in Eqs. (1-3) are all approximately infinitesimal, say $\leq h \ll 1$, in magnitude. The sampling period ϵ is such that $h^2\ll\epsilon\ll h$. All other quantites are of order unity, including the components of B^{-1}, R^{-1}, and P_0^{-1}. It is assumed that the corresponding standard linear-quadratic-Gaussian problem (i.e., with all three-way matrices zero) is sufficiently observable and controllable that the prior expected value of x^Tx remains of order unity under the optimal control law. Also, control laws for which the control is not always of order unity (except perhaps for a negligible set of realizations) are considered inadmissible here.

It is helpful to define the parameters

$$\Gamma(t)=\tfrac{1}{2}[Q(t)\bar{\Gamma}^T(t)+\bar{\Gamma}(t)Q(t)]$$
$$\Psi(t)=\tfrac{1}{2}[Q(t)\bar{\Psi}^T(t)+\bar{\Psi}(t)Q(t)]$$
$$\Omega(t)=\tfrac{1}{2}[R(t)\bar{\Omega}^T(t)+\bar{\Omega}(t)R(t)]$$
$$\Upsilon(t)=\tfrac{1}{2}[R(t)\bar{\Upsilon}^T(t)+\bar{\Upsilon}(t)R(t)]$$

and to note that the state measurements here are equivalent to a step-function process z such that

$$z(t)=\frac{y(i\epsilon)-y[(i-1)\epsilon]}{\epsilon}\quad \text{for}\ t\ \text{in}\ [i\epsilon,(i+1)\epsilon) \tag{4}$$

which constitute measurements corrupted by "almost white" Gaussian noise.

Received April 10, 1978; Synoptic received Oct. 5, 1978. Full paper available from National Technical Information Service, Springfield, Va. 22151, as N78-33820 at the standard price (available upon request).

Index categories: Guidance and Control; Analytical and Numerical Methods.

*Operations Research Analyst, Space Systems Division. Member AIAA.

A key step in deriving the results is to approximate the conditional probability density of the current state given the available measurements by a first-order Edgeworth expansion of the form

$$p(x) \cong \left[\frac{\exp[-\tfrac{1}{2}(x-\hat{x})^{\mathrm{T}} V^{-1}(x-\hat{x})]}{(2\pi)^{\frac{1}{2}n}|V|^{\frac{1}{2}}}\right]$$

$$[1+(x-\hat{x})^{\mathrm{T}}\mathrm{Tr}\{(V^{-1}\Lambda V^{-1})' V^{-1}[\tfrac{1}{3}(x-\hat{x})(x-\hat{x})^{\mathrm{T}}-V]\}]$$

where Λ is a symmetric three-way matrix of order h and n is the dimension of x. This class of expansions is formally "self-reproducing" here as an approximation of the conditional state density to first order in h. The parameters of this approximation evolve according to the equations

$$V=P+2D \tag{5}$$

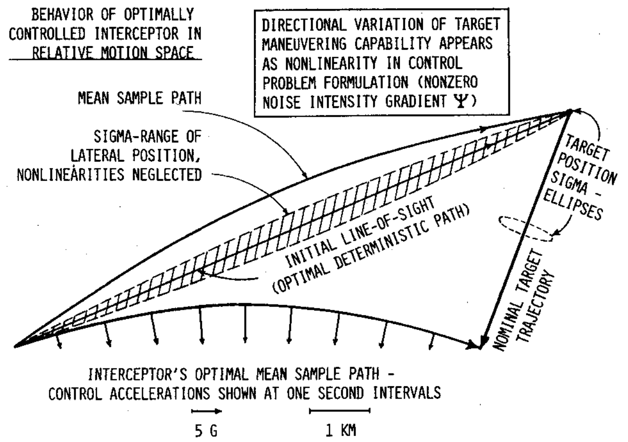

Fig. 1 Planar free-space intercept problem.

$$\dot{P}=FP+PF^{\mathrm{T}}+Q-P[H^{\mathrm{T}}R^{-1}H+(2/\epsilon)\,\mathrm{Tr}(R^{-1}\Upsilon R^{-1}\Upsilon)]P \qquad P(0)=P_0 \tag{6}$$

$$\dot{\Lambda}=\nabla+\nabla'+\nabla''$$

where

$$\nabla=(F-PH^{\mathrm{T}}R^{-1}H)\Lambda+\Psi'P+P\Delta P-(P\Phi P)'R^{-1}HP+(PH^{\mathrm{T}}R^{-1}\Upsilon R^{-1}HP)'P \qquad \Lambda(0)=0 \tag{7}$$

$$\dot{D}=(F-PH^{\mathrm{T}}R^{-1}H)D+D(F^{\mathrm{T}}-H^{\mathrm{T}}R^{-1}HP)+(\Gamma+P\Theta'+\Theta^{\mathrm{T}}P+PH^{\mathrm{T}}R^{-1}\Omega R^{-1}HP)'u+[\Psi+P\Delta'+\Delta''$$
$$P+P(H^{\mathrm{T}}R^{-1}\Upsilon R^{-1}H-H^{\mathrm{T}}R^{-1}\Phi''-\Phi'R^{-1}HP]'\hat{x}+\{\Lambda H^{\mathrm{T}}+[P(\Phi-H^{\mathrm{T}}R^{-1}\Upsilon'-\Upsilon''R^{-1}H)P]'\}R^{-1}(z-H\hat{x}) \qquad D(0)=0 \tag{8}$$

$$\dot{\hat{x}}=(F+\Theta'u)\hat{x}+(G+\Theta^{\mathrm{T}}\hat{x})u+\mathrm{Tr}[\Delta(\hat{x}\hat{x}^{\mathrm{T}}+P)]+PH^{\mathrm{T}}R^{-1}\{z-H\hat{x}-\mathrm{Tr}[\Phi(\hat{x}\hat{x}^{\mathrm{T}}+P)]\}+2[DH^{\mathrm{T}}+P\Phi\hat{x}$$
$$-PH^{\mathrm{T}}R^{-1}(\Upsilon'\hat{x}+\Omega'u)]R^{-1}(z-H\hat{x})+\tfrac{1}{2}P\mathrm{Tr}\{[\bar{\Upsilon}^{\mathrm{T}}(R+2\Upsilon'\hat{x}+2\Omega'u)^{-1}+(R+2\Upsilon'\hat{x}+2\Omega'u)^{-1}\bar{\Upsilon}][(z-H\hat{x})(z-H\hat{x})^{\mathrm{T}}$$
$$-HPH^{\mathrm{T}}-(1/\epsilon)(R+2\Upsilon'\hat{x}+2\Omega'u)]\} \qquad \hat{x}(0)=\hat{x}_0 \tag{9}$$

where the "derivative" $\dot{\hat{x}}(t)$ should be interpreted as $(1/\epsilon)\{\hat{x}(i\epsilon)-\hat{x}[(i-1)\epsilon]\}$ for $t=i\epsilon$, $i=1,2,\ldots$, with $\hat{x}(\tau)=\hat{x}(i\epsilon)$ for *all* τ in $[i\epsilon,\ (i+1)\epsilon)$, in order to avoid the difficulties described by Wong and Zakai.[4] A standard dynamic programming argument using $\hat{x}$ *and* D as state variables then gives the optimal control law and corresponding performance formally to first order in h as

$$u=-B^{-1}\{G^{\mathrm{T}}S\hat{x}+\mathrm{Tr}[(SGB^{-1}\Xi B^{-1}G^{\mathrm{T}}S+\Theta S+S\Theta^{\mathrm{T}}+\Pi G)\hat{x}\hat{x}^{\mathrm{T}}]+G^{\mathrm{T}}\phi+c+\mathrm{Tr}[(S+Y)(\Gamma+P\Theta+\Theta^{\mathrm{T}}P)+PH^{\mathrm{T}}R^{-1}\Omega R^{-1}HPY]\} \tag{10}$$

$$J=\tfrac{1}{2}\mathrm{Tr}[S_0(\hat{x}_0\hat{x}_0^{\mathrm{T}}+P_0)+\int_0^{t_f}(SQ+SGB^{-1}G^{\mathrm{T}}SP)\,\mathrm{d}t]+\phi_0^{\mathrm{T}}\hat{x}_0+\tfrac{1}{3}\hat{x}^{\mathrm{T}}{}_0\mathrm{Tr}(\Pi_0\hat{x}_0\hat{x}_0^{\mathrm{T}}) \tag{11}$$

where

$$\dot{S}=-SF-F^{\mathrm{T}}S-A+SGB^{-1}G^{\mathrm{T}}S \qquad S(t_f)=S_f \tag{12}$$

$$\dot{Y}=Y(PH^{\mathrm{T}}R^{-1}H-F)+(H^{\mathrm{T}}R^{-1}HP-F^{\mathrm{T}})Y-SGB^{-1}G^{\mathrm{T}}S \qquad Y(t_f)=0 \tag{13}$$

$$\dot{\Pi}=(SGB^{-1}\Xi B^{-1}G^{\mathrm{T}}S)'B^{-1}G^{\mathrm{T}}\overset{*}{S}-\Sigma+\nabla+\nabla'+\nabla''$$

where

$$\nabla=\Pi(GB^{-1}G^{\mathrm{T}}S-F)-\Delta'S+S\Theta'B^{-1}G^{\mathrm{T}}S+SGB^{-1}(\Theta')^{\mathrm{T}}S \qquad \Pi(t_f)=\Pi_f \tag{14}$$

$$\dot{\phi}=(SGB^{-1}G^{\mathrm{T}}-F^{\mathrm{T}})\phi+c+SGB^{-1}\mathrm{Tr}[(S+Y)(\Gamma+P\Theta+\Theta^{\mathrm{T}}P)+PH^{\mathrm{T}}R^{-1}\Omega R^{-1}HPY]-S\,\mathrm{Tr}(P\Delta)-\mathrm{Tr}[P\Sigma+PH^{\mathrm{T}}R^{-1}HP\Pi$$
$$+P(H^{\mathrm{T}}R^{-1}\Upsilon R^{-1}H-H^{\mathrm{T}}R^{-1}\Phi''-\Phi'R^{-1}H)PY+(S+Y)(\Psi+P\Delta'+\Delta''P)] \qquad \Phi(t_f)=\mathrm{Tr}[\Pi_f P(t_f)] \tag{15}$$

If $\bar{\Upsilon}=0$, the value of ϵ (which actually may be time-varying in the preceding) is unimportant as long as $\epsilon \ll h$; in fact, y itself can be taken as the measurement process. If x can be measured exactly, state estimation is superfluous, and a similar dynamic programming argument gives the optimal control law and performance formally to first order as Eqs. (10-15) with $\hat{x}$ replaced by x, $PH^{\mathrm{T}}R^{-1}HP$ by Q, and P and Y otherwise by zero.

Figure 1 shows the significance these effects of nonlinearities can attain under favorable but realistic conditions. The parameter values are characteristic of an air-to-air homing missile.[1] The interceptor can accelerate perpendicular to its flight path and receives noisy angular line-of-sight measurements to a target with white-noise acceleration of directionally varying intensity. This directional variation causes the optimal control law's mean sample path in relative motion space, which would otherwise be the initial line-of-sight, to assume the form shown to first order in the noise intensity gradients. In these computations, the intended initial interceptor heading was also optimized by minimizing the performance given by Eq. (11).

References

[1] Bryson, A.E. and Ho, Y.-C., *Applied Optimal Control*, Hemisphere, Wash., 1975, Chap. 14.

[2] Gustafson, D.E. and Speyer, J.L., "Design of Linear Regulators for Nonlinear Stochastic Systems," *Journal of Spacecraft and Rockets*, Vol. 12, June 1975, pp. 351-358.

[3] Alspach, D.L., "A Stochastic Control Algorithm for Systems with Control Dependent Plant and Measurement Noise," *Computers and Electrical Engineering*, Vol. 2, Nov. 1975, pp. 297-306.

[4] Wong, E. and Zakai, M., "On the Relation Between Ordinary and Stochastic Differential Equations," *International Journal of Engineering Science*, Vol. 3, July 1965, pp. 213-229.