Active Dual Control for Stochastic Linear Systems: Algorithm

Vinay A. Bavdekar[a], Ali Mesbah[a]

[a] *Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA*

1. Optimal-Cost-to-Go Computations

1.1. Problem statement

Consider the following discrete-time linear system model

$$x_{k+1} = A(\theta_k, k)\, x_k + B(\theta_k, k)\, u_k + w_k, \tag{1a}$$

$$y_k = C(\theta_k, k)\, x_k + v_k, \tag{1b}$$

where $x \in \mathbb{R}^n$ represents the system states; $u_k \in \mathbb{R}^1$ represents the manipulated inputs; $y \in \mathbb{R}^r$ represents the measurements; $w \sim \mathcal{N}(0, Q) \in \mathbb{R}^n$ represents the stochastic disturbances in the system dynamics that possess a Gaussian distribution with mean 0 and covariance Q; and $v \sim \mathcal{N}(0, R) \in \mathbb{R}^r$ is the measurement noise with mean 0 and covariance R. The parameters $\theta \in \mathbb{R}^p$ are the uncertain parameters governed by a Markov process satisfying

$$\theta_{k+1} = D_k \theta_k + w_{\theta,k}, \tag{2}$$

where D_k is a known matrix; and $w_\theta \sim \mathcal{N}(0, Q_p) \in \mathbb{R}^p$ represents the noise in the parameter dynamics. The initial conditions of the system model are

Email addresses: `bavdekar@berkeley.edu` (Vinay A. Bavdekar), `mesbah@berkeley.edu` (Ali Mesbah)

defined as

$$x_0 \sim \mathcal{N}\left(\hat{x}_0, P_0\right), \quad \theta_0 \sim \mathcal{N}\left(\hat{\theta}_0, P_{\theta,0}\right).$$

It is assumed that the parameter θ_k enters linearly in $A(\cdot, k)$, $B(\cdot, k)$ and $C(\cdot, k)$.

The objective is to find an admissible control sequence $U_{N-1} \triangleq [u_1, \cdots, u_{N-1}]$ that minimizes the following cost functional

$$J(U) = \frac{1}{2} E\left\{ \left[x_N - x_{SP}\right]^\top W_N \left[x_N - x_{SP}\right] + \sum_{k=0}^{N-1} \left[x_k - x_{SP}\right]^\top W_k \left[x_k - x_{SP}\right] + \lambda_k u_k^2 \right\}, \tag{3}$$

where N is the horizon over which the control actions are implemented; $W_j > 0,\ \{j = 1, \ldots, N\}$ is the weighting matrix; x_{SP} is the set-point; and λ is the weight on the input. If u_k is a vector, i.e. $u_k \in \mathbb{R}^m,\ m > 1$, then λ is a diagonal matrix.

1.2. The optimal-cost-to-go

Let k denote the present time. Denote the augmented state as $z_{o,k} = \left[x_{o,k}^\top \;\; \theta_{o,k}^\top\right]^\top$, where the subscript o refers to the nominal value. The nominal control sequence is denoted $\{u_{0,j}\}_{j=k+1}^{N}$. In this case, the nominal sequence is obtained using the certainty equivalence control objective. The nominal trajectory of the augmented state space is generated as follows

$$z_{o,j+1} \triangleq \begin{bmatrix} x_{0,j+1} \\ \theta_{o,j+1} \end{bmatrix} = f_{o,j} \triangleq \begin{bmatrix} f_{o,j}^x \\ f_{o,j}^\theta \end{bmatrix} \triangleq \begin{bmatrix} A_{o,j} x_{o,j} + B_{o,j} u_{o,j} \\ D_{o,j} \theta_{o,j} \end{bmatrix}, \tag{4}$$

where the superscripts denote matrix partitions. The initial condition is $z_{o,k} = \hat{z}_{k|k}$, which can be obtained from the EKF. The covariance of $z_{o,k}$ is partitioned as

$$P_{z,k} = \begin{bmatrix} P_{x,k} & P_{x\theta,k} \\ P_{x\theta,k}^\top & P_{\theta,k} \end{bmatrix} \tag{5}$$

In order to analyze the dual effect of the control actions, a second order perturbation analysis of the control objective is carried out over the horizon of the control actions. Define the Jacobian of the augmented space, evaluated at $z_{o,j} = [x_{o,j} \quad \theta_{o,j}]^\top$, as follows

$$F_{o,j} \triangleq \frac{\partial f_{o,j}}{\partial z_{o,j}} = \begin{bmatrix} F_{o,j}^{xx} & F_{o,j}^{x\theta} \\ F_{o,j}^{\theta x} & F_{o,j}^{\theta\theta} \end{bmatrix}. \tag{6}$$

The optimal-cost-to-go is now computed as follows

1. Obtain $\hat{x}_{k+1|k}$, $\hat{\theta}_{k+1|k}$, and $\mathsf{P}_{z,k+1|k}$ as follows (Eqs. 4.39–4.41)

$$\hat{x}_{k+1|k} = A(\hat{\theta}_{k|k}, k)\hat{x}_{k|k} + B(\hat{\theta}_{k|k}, k)u_k + \frac{1}{2}\sum_{i=1}^{n} e^{(i)}\mathrm{tr}\left(F_{k|k}^{xx,(i)} P_{x,k|k}\right), \tag{7}$$

$$\hat{\theta}_{k+1|k} = D_k \hat{\theta}_{k|k}, \tag{8}$$

$$P_{z,k+1|k} = F_{o,k} P_{z,k|k} F_{o,k}^\top + Q_z + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} e^{(i)} \left(e^{(j)}\right)^\top \mathrm{tr}\left(F_{o,k}^{(i)} P_{z,k|k} F_{o,k}^{(j)} P_{z,k|k}\right), \tag{9}$$

where $e^{(i)}$, $\forall\, i = 1, \dots, n$ represents the standard Euclidean basis, and

$$Q_z = \begin{bmatrix} Q_x & 0 \\ 0 & Q_\theta \end{bmatrix}. \tag{10}$$

2. Generate the predictions of the parameters as

$$\hat{\theta}_{0,j+1} = D_j \hat{\theta}_{o,j} \quad \forall\, j = k, \dots, N-1 \tag{11}$$

3. Back calculate the following matrices (Eq. 4.30–4.31)

$$K_{o,j} = A_{o,j}^{\top} \left[I - \mu_{o,j} K_{o,j+1} B_{o,j} B_{o,j}^{\top} \right] K_{o,j+1} A_{o,j} + W_j; \quad K_{o,N} = W_N, \tag{12}$$

$$p_{o,j} = A_{o,j}^{\top} \left[I - \mu_{o,j} K_{o,j+1} B_{o,j} B_{o,j}^{\top} \right] p_{o,j+1} - W_j x_{SP,j}; \quad p_{o,N} = -W_N x_{SP,N}, \tag{13}$$

where (Eq. 4.29)

$$\mu_{o,j} = \left[\lambda_j + B_{o,j}^{\top} K_{o,j+1} B_{o,j} \right]^{-1}. \tag{14}$$

4. Compute the optimal control inputs as follows (Eq. 4.28)

$$u_{o,j}^{*} = -\mu_{o,j} B_{o,j}^{\top} \left[K_{o,j+1} A_{o,j} x_{o,j} + p_{o,j+1} \right], \ \forall \ j = k+1, \ldots, N, \tag{15}$$

where $x_{o,j}$, $\forall \ j = k+1, \ldots, N$, represents the nominal predicted trajectory of the states.

5. Back compute the following matrices (Eqs. 4.16–4.17)

$$K^{\theta x}_{o,j} = \left[\left(F^{x\theta}_{o,j}\right)^\top K_{o,j+1} + D_j^\top K^{\theta x}_{o,j+1}\right] A_{o,j} \tag{16}$$

$$- \mu_{o,j} \left\{ \left[\left(F^{x\theta}_{o,j}\right)^\top K_{o,j+1} + D_j^\top K^{\theta x}_{o,j+1}\right] B_{o,j} + \left[\sum_{i=1}^{n} \left(e^{(i)}\right)^\top p^x_{o,j+1} B^{(i)}_{\theta,j}\right]^\top \right\} \tag{17}$$

$$\cdot \left\{B^\top_{o,j} K_{o,j+1} A_{o,j}\right\}; \quad K^{\theta x}_{o,N} = 0, \tag{18}$$

$$K^{\theta\theta}_{o,j} = \left(F^{x\theta}_{o,j}\right)^\top K_{o,j+1}\left(F^{x\theta}_{o,j}\right) + D_j^\top K^{\theta x}_{o,j+1}\left(F^{x\theta}_{o,j}\right) + \left(F^{x\theta}_{o,j}\right)^\top K^{x\theta}_{o,j+1} D_j + D_j^\top K^{\theta\theta}_{o,j+1} D_j \tag{19}$$

$$- \mu_{o,j} \left\{ B_j^\top \left[K_{i,j+1} F^{x\theta}_j + K^{x\theta}_{o,j+1} D_j\right] + \sum_{i=1}^{n} \left(e^{(i)}\right)^\top p^x_{o,j+1} B^{(i)}_{\theta,j} \right\}^\top \tag{20}$$

$$\cdot \left\{ B_j^\top \left[K_{i,j+1} F^{x\theta}_j + K^{x\theta}_{o,j+1} D_j\right] + \sum_{i=1}^{n} \left(e^{(i)}\right)^\top p^x_{o,j+1} B^{(i)}_{\theta,j} \right\}; \quad K^{\theta\theta}_{o,N} = 0, \tag{21}$$

$$p^x_{o,j} = K_{o,j} x_{o,j} + p_{o,j}, \tag{22}$$

where $B^{(i)}_{\theta,j}$ represents the jacobian of the input-to-state function with respect to θ.

6. Form the matrix (Eq. 4.34)

$$\bar{K}_j = \begin{bmatrix} K_{o,j} & \left(K^{\theta x}_{o,j}\right)^\top \\ K^{\theta x}_{o,j} & K^{\theta\theta}_{o,j} \end{bmatrix} \tag{23}$$

7. Generate the predictions of the mean and covariance of the nominal

trajectory as follows (Eqs. 4.26, 4.28, 4.21–4.24)

$$x_{o,j+1|k} = A\left(\hat{\theta}_{o,j}, j\right) x_{o,j} + B\left(\hat{\theta}_{o,j}, j\right) u^*_{o,j},\ \forall\ j = k+1, \ldots, N, \tag{24a}$$

$$V_{o,j+1} = P_{z,j+1|j} C_z^\top \left[C_z P_{z,j+1|j} C_z^\top + R\right]^{-1},\ \forall\ j = k, \ldots, N-1 \tag{24b}$$

$$P_{z,j+1|j+1} = \left[I - V_{o,j+1} C_z\right] P_{z,j+1|j},\ \forall\ j = k, \ldots, N-1, \tag{24c}$$

$$P_{z,j+1|j} = f_{o,j} P_{z,j|j} f_{o,j}^\top + Q_z,\ \forall\ j = k+1, \ldots, N, \tag{24d}$$

where $C_z = \begin{bmatrix} \frac{\partial C(\hat{\theta}_j, j)}{\partial x} & \frac{\partial C(\hat{\theta}_j, j)}{\partial \theta} \end{bmatrix}$.

8. Compute the one-step dual cost-to-go as follows (Eq. 4.42)

$$\begin{aligned} J_d\left[u_k\right] =& \frac{1}{2}\lambda_k u_k^2 + \frac{1}{2}\hat{x}_{k+1}^\top K_{o,k+1} \hat{x}_{k+1|k}^\top + p_{o,k+1}^\top \hat{x}_{k+1|k} \\ & + \frac{1}{2}\mathrm{tr}\left\{ \sum_{j=k+1}^{N-1} W_j P_{x,j|j} + \left[P_{z,k+1|k} - P_{z,k+1|k+1}\right] \bar{K}_{k+1} \right. \\ & \left. \sum_{j=k+1}^{N-1} \left[P_{z,j+1|j} - P_{z,j+1|j+1}\right] \bar{K}_{j+1} \right\} \end{aligned} \tag{24e}$$