УДК 531.36:62–50

© 1996 г. С.В. Соколов

О РЕШЕНИИ ПРОБЛЕМЫ СИНТЕЗА СТОХАСТИЧЕСКОГО ОПТИМАЛЬНОГО УПРАВЛЕНИЯ НА ОСНОВЕ НЕЛИНЕЙНЫХ ВЕРОЯТНОСТНЫХ КРИТЕРИЕВ

Рассмотрен метод, позволяющий осуществлять точный синтез законов управления нелинейными стохастическими объектами, оптимальных в смысле вероятностных критериев общего вида. Показаны преимущества предложенного метода по сравнению с традиционными, рассмотрен пример его практического использования.

Существующие методы синтеза стохастического управления, оптимального в смысле математического ожидания некоторого заданного функционала [1, 2], не позволяют использовать в качестве оптимизируемых вероятностных критериев более общие формы, нелинейно зависящие от плотности распределения параметров состояния объекта, например критерий минимума энтропии вектора состояния, критерий Кульбака и т.д.

В связи с этим возникает проблема разработки такого подхода к синтезу оптимального управления стохастическими объектами, который позволял бы построить точный закон управления при самой общей нелинейной зависимости критерия оптимизации от плотности распределения вектора состояния.

1. Постановка задачи. Пусть стохастический объект описывается нелинейным векторным дифференциальным уравнением размерности N в симметризованной форме

$$\mathbf{X}^{\cdot} = \mathbf{f}(\mathbf{X},t) + f_0(\mathbf{X},t)\mathbf{V}_t + \mathbf{U}(\mathbf{X},t) \tag{1.1}$$

где $\mathbf{f} = (f_1, f_2,\dots, f_N), f_0 = \| f_{0_{ik}} \|$ – известные нелинейные векторная и матричная функции, $\mathbf{V}_t$ – белый гауссовский нормированный вектор-шум размерности N_1, $\mathbf{U} = (U_1, U_2,\dots, U_N)$ – искомый вектор управления.

На временном интервале $T = [t_0, t_k]$ требуется определить вектор управления $\mathbf{U}(\mathbf{X}, t)$ из условия минимума вероятностного функционала J, заданного на ограниченном множестве $\mathbf{X}_* \in [\mathbf{X}_{\min}, \mathbf{X}_{\max}]$ и нелинейно зависящего как от управления $\mathbf{U}(\mathbf{X}, t)$, так и от функции плотности вероятности $\rho(\mathbf{X}, t)$ процесса $\mathbf{X}_t$:

$$J = \int_T \int_{\mathbf{X}_*} \Phi[\rho(\mathbf{X},t), \mathbf{U}(\mathbf{X},t)]\, d\mathbf{X}dt = \int_T W dt \tag{1.2}$$

где Φ – известная нелинейная функция, учитывающая в общем случае возможные аналитические ограничения на вектор управления.

Так как входящая в функционал (1.2) и подлежащая управлению плотность $\rho(\mathbf{X}, t)$ процесса $\mathbf{X}_t$ описывается уравнением Фоккера–Планка–Колмогорова

$$\frac{\partial\rho(\mathbf{X},t)}{\partial t} = L\{\mathbf{a}, b, \rho(\mathbf{X},t)\} \tag{1.3}$$
$$\mathbf{a} = (a_1, a_2,\dots, a_N), \quad b = \| b_{ij} \|, \quad i,j = 1,2,\dots,N$$

$$L\{\mathbf{a},b,\rho(\mathbf{X},t)\} = -\sum_i \frac{\partial}{\partial X_i}\{a_i(\mathbf{X},t)\rho(\mathbf{X},t)\} + \frac{1}{2}\sum_i\sum_j \frac{\partial^2}{\partial x_i \partial x_j}\{b_{ij}(\mathbf{X},t)\rho(\mathbf{X},t)\}$$

$$a_i(\mathbf{X},t) = f_i(\mathbf{X},t) + U_i(\mathbf{X},t) + \frac{1}{4}\sum_k\sum_j f_{0_{jk}}(\mathbf{X},t)\frac{\partial}{\partial x_j} f_{0_{ik}}(\mathbf{X},t) =$$

$$= u_i(\mathbf{X},t) + a_{1i}(\mathbf{X},t), \quad b_{ij}(\mathbf{X},t) = \frac{1}{2}\sum_k f_{0_{ik}}(\mathbf{X},t)\, f_{0_{jk}}(\mathbf{X},t)$$

являющимся уравнением в частных производных, то для дальнейшего решения поставленной задачи необходимо привлечение методов теории оптимального управления системами с распределенными параметрами. Здесь и всюду далее суммирование по i, j ведется от 1 до N, по k – от 1 до N_1.

2. Синтез оптимального управления. Оптимальное управление будем искать в классе ограниченных непрерывных функций со значениями из открытой области $\mathbf{U}_*$. Для его построения используем метод динамического программирования, согласно которому задача сводится к решению функционального уравнения [3]

$$\min_{\mathbf{U}\in\mathbf{U}_*}\left\{\frac{dV}{dt} + W\right\} = 0 \qquad (2.1)$$

при конечном условии $V(t_k) = 0$ относительно оптимального функционала V, параметрически зависящего от времени $t \in T$ и определенного на множестве функций ρ, удовлетворяющих уравнению (1.3).

Для линейных систем функционал отыскивается в виде интегральной квадратичной формы [3]

$$V = \int_{\mathbf{X}} \upsilon(\mathbf{X},t)\rho^2(\mathbf{X},t)d\mathbf{X}$$

откуда имеем

$$\frac{dV}{dt} + W = \int_{\mathbf{X}}\left\{\frac{d\upsilon}{dt}\rho^2 + 2\upsilon\rho L\{\mathbf{a}_1,b,\rho\} + \Phi[\rho,\mathbf{U}] - 2\upsilon\rho\sum_i\left(\frac{\partial u_i}{\partial x_i}\rho + u_i\frac{\partial\rho}{\partial x_i}\right)\right\}d\mathbf{X} \qquad (2.2)$$

$$\mathbf{a}_1 = (a_{11}, a_{12}, \ldots, a_{1N})$$

Анализ полученного выражения показывает, что определение вектора $\mathbf{U}(\mathbf{X}, t)$ из решения функционального уравнения (2.1) сводится к классической задаче отыскания вектор-функции, реализующей минимум определенного интеграла (2.2). При этом вектор-функция $\mathbf{U}(\mathbf{X}, t)$, являющаяся решением данной задачи, должна удовлетворять системе уравнений Эйлера

$$-\frac{\partial}{\partial x_i}[2\upsilon\rho^2] - \frac{\partial\Phi}{\partial u_i} + 2\upsilon\rho\frac{\partial\rho}{\partial x_i} = 0, \quad i = 1,2,\ldots,N$$

или

$$\left(\frac{\partial\Phi}{\partial\mathbf{U}}\right)^T = -\rho\frac{\partial}{\partial\mathbf{X}}[2\upsilon\rho]^T \qquad (2.3)$$

В общем случае (2.3) представляет собой систему нелинейных уравнений относительно компонентов вектора управления, допускающую аналитическое разрешение только в некоторых частных случаях, например, когда функция $\Phi[\rho, \mathbf{U}]$ имеет следующий вид:

$$\Phi = \Phi_0[\rho] + \sum_i \psi_i(u_i)$$

где ψ_i – аналитические функции, первые производные которых допускают обращение, или

$$\Phi = \Phi_0[\rho] + [\mathbf{U}(\mathbf{X},t) - \mathbf{U}_0(\mathbf{X},t)]^T D_U(\mathbf{X},t)[\mathbf{U}(\mathbf{X},t) - \mathbf{U}_0(\mathbf{X},t)]$$

где $D_U(\mathbf{X}, t)$ – известная квадратная симметричная матрица, $\mathbf{U}_0(\mathbf{X}, t)$ – известный вектор.

В последнем случае использование квадратичной формы вектора $\mathbf{U}$ обусловливается, как правило, требованием обеспечения минимального отклонения формы искомого управления от заданного вектора $\mathbf{U}_0$, определяемого в свою очередь возможностями технической реализации управления. При этом уравнение (2.3) разрешается относительно $\mathbf{U}$ следующим образом:

$$\mathbf{U}_{\text{opt}} = \mathbf{U}_0 - \rho D_U^{-1}\left[\frac{\partial}{\partial \mathbf{X}}(2\nu\rho)\right]^T \tag{2.4}$$

Подстановка найденного закона оптимального управления в (1.3) и (2.2) (при учете условия $dV/dt + W = 0$, формируемого при оптимальном управлении $\mathbf{U}_{\text{opt}}$ [3]) позволяет построить систему уравнений в частных производных

$$\begin{gathered}
\frac{\partial \rho}{\partial t} = L\{\mathbf{a}_1, b, \rho\} - A(\rho,\nu), \quad \rho(\mathbf{X},t_0) = \rho_0 \\
A(\rho,\nu) = \sum_i \frac{\partial}{\partial x_i}\left\{\left(\mathbf{U}_0 - \rho D_U^{-1}\left[\frac{\partial}{\partial \mathbf{X}}(2\nu\rho)\right]^T\right)_{(i)}\rho\right\} \\
\frac{\partial \nu}{\partial t} = -2\nu\rho^{-1}L\{\mathbf{a}_1, b, \rho\} - \Phi_0[\rho]\rho^{-2} - \left[\frac{\partial}{\partial \mathbf{X}}(2\nu\rho)\right]D_U^{-1}\left[\frac{\partial}{\partial \mathbf{X}}(2\nu\rho)\right]^T + 2\nu\rho^{-1}A(\rho,\nu) \\
\nu(\mathbf{X},t_k) = 0
\end{gathered} \tag{2.5}$$

Индекс i введен для обозначения i-го компонента вектора.

Решение этой системы исчерпывает по существу решение поставленной проблемы.

3. Анализ путей вычислительной реализации оптимального управления. С точки зрения точности и сложности формирования оптимального управления интересен сравнительный анализ решения системы (2.5) и системы сопряженных уравнений, получаемых при использовании традиционного подхода [1, 2]. Если в случае (2.5) оптимальное управление зависит от плотности ρ пропорционально-дифференциально, то управление, полученное традиционно, зависит от ρ интегрально, что приводит для N-мерного вектора состояния к необходимости совместного интегрирования $(2N + 1)$-мерной системы сопряженных обыкновенных интегродифференциальных уравнений и интегродифференциального уравнения в частных производных для N-мерной функции ρ, что существенно сложнее решения системы для уравнений в частных производных (2.5). Тем не менее от этого сложность решения системы (2.5), общих методов точного аналитического решения которой в настоящее время не существует, не уменьшается. Не останавливаясь на многочисленных приближенных методах решения данной задачи, ориентированных на компромисс между необходимой точностью и объемом вычислительных затрат, в качестве одного из методов решения этой проблемы рассмотрим далее метод, основанный на разложении функций ν и ρ в ряд по некоторой системе ортонормированных функций векторного аргумента:

$$\nu(\mathbf{X},t) = \sum_\mu \alpha_\mu(t)\varphi_\mu(\mathbf{X}) = \boldsymbol{\varphi}^T\boldsymbol{\alpha}$$

$$\rho(\mathbf{X},t) = \sum_\mu \beta_\mu(t)\varphi_\mu(\mathbf{X}) = \boldsymbol{\varphi}^T\boldsymbol{\beta}$$

где μ – индекс, пробегающий множество значений от $(0,..., 0)$ до $(N_2,..., N_2)$ [3], φ – вектор ортонормированных функций аргумента $\mathbf{X}$, α,β – векторы коэффициентов соответствующих разложений.

В этом случае решение сводится к решению двухточечной краевой задачи интегрирования системы следующих обыкновенных дифференциальных уравнений:

$$\beta^{\cdot} = \int\limits_{\mathbf{X}} \varphi L[\mathbf{a}_1, b, \varphi^T \beta] d\mathbf{X} - \int\limits_{\mathbf{X}} \varphi A(\varphi^T \beta, \varphi^T \alpha) d\mathbf{X}$$

$$\alpha^{\cdot} = \int\limits_{\mathbf{X}} \varphi \Big\{ -2\varphi^T \alpha (\varphi^T \beta)^{-1} L[\mathbf{a}_1, b, \varphi^T \beta] - \Phi_0[\varphi^T \beta](\varphi^T \beta)^{-2} -$$

$$-\left[\frac{\partial}{\partial \mathbf{X}}(2\varphi^T \alpha \varphi^T \beta)\right] D_U^{-1} \left[\frac{\partial}{\partial \mathbf{X}}(2\varphi^T \alpha \varphi^T \beta)\right]^T + 2\varphi^T \alpha (\varphi^T \beta)^{-1} A(\varphi^T \beta, \varphi^T \alpha) \Big\} d\mathbf{X} \qquad (3.1)$$

при краевых условиях $\alpha(t_k) = 0$, $\beta(t_0) = \beta_0$, причем значения компонентов β_0 определяются из разложения функции $\rho(\mathbf{X}, t_0) = \rho_0$.

С точки зрения практической реализации интегрирование системы (3.1) при краевых условиях оказывается проще, чем интегрирование (2.5), но с точки зрения организации процесса оценивания в реальном времени по-прежнему остается весьма трудновыполнимой задачей. Более того, целесообразность такого прямого подхода весьма сомнительна по следующим причинам. Во-первых, оказывается велик объем необходимых временных и вычислительных затрат, во-вторых, исключается возможность настройки вектора $\mathbf{U}$ в реальном времени, и в-третьих, в процессе приборной реализации, как правило, все равно не удается выдержать точно заданные значения $\mathbf{U}$.

Таким образом, в данном случае вполне обосновано использование приближенных методов решения задачи (3.1), в качестве одного из которых рассмотрим далее метод инвариантного погружения [4], обеспечивающий искомое приближенное решение в реальном масштабе времени.

Так как применение данного метода предполагает задание всех составляющих искомого приближенно оцениваемого вектора в дифференциальной форме, то для реализации возможности синтеза вектора $\mathbf{U}$ с помощью данного метода в реальном времени, введем фиктивную переменную ϑ, позволяющую в дальнейшем учесть выражение (2.4) в виде дифференциального уравнения

$$\vartheta^{\cdot} = \mathbf{U}_{\text{opt}}(\varphi^T \alpha,\ \varphi^T \beta)$$

образующего с уравнениями (3.1) единую систему. Применение метода инвариантного погружения приводит в этом случае к следующей системе уравнений:

$$\begin{vmatrix} \hat{\vartheta}^{\cdot} \\ \hat{\beta}^{\cdot} \end{vmatrix} = \begin{vmatrix} \mathbf{U}_0 \\ \int\limits_{\mathbf{X}} \varphi B(\mathbf{a}_1, b, \varphi^T \hat{\beta}, \mathbf{U}_0) d\mathbf{X} \end{vmatrix} - D \int\limits_{\mathbf{X}} \varphi \Phi_0[\varphi^T \hat{\beta}](\varphi^T \hat{\beta})^{-2} d\mathbf{X}$$

$$D^{\cdot} = 2\int\limits_{\mathbf{X}} \varphi \left\{ \frac{\partial}{\partial \hat{\beta}} B(\mathbf{a}_1, b, \varphi^T \hat{\beta}, \mathbf{U}_0) \right\} d\mathbf{X} \cdot D + D \int\limits_{\mathbf{X}} \varphi \{ 2\varphi^T (\varphi^T \hat{\beta})^{-1} B(\mathbf{a}_1, b, \varphi^T \hat{\beta}, \mathbf{U}_0) \} d\mathbf{X} +$$

$$+\int\limits_{\mathbf{X}} \varphi \left\{ \sum_i \frac{\partial}{\partial x_i} \left(\left[(\varphi^T \hat{\beta}) D_U^{-1} \frac{\partial}{\partial \mathbf{X}} (2\varphi^T [\varphi^T \hat{\beta}]) \right]_{(i)} \varphi^T \hat{\beta} \right) \right\} d\mathbf{X} -$$

$$-2D \int\limits_{\mathbf{X}} \varphi \left\{ \frac{\partial}{\partial \hat{\beta}} [\Phi_0[\varphi^T \hat{\beta}](\varphi^T \hat{\beta})^{-2}] \right\} d\mathbf{X} \cdot D$$

$$B(\mathbf{a}_1, b, \varphi^T \hat{\beta}, \mathbf{U}_0) = L[\mathbf{a}_1, b, \varphi^T \hat{\beta}] - \sum_i \frac{\partial}{\partial x_i} [u_{0(i)} \varphi^T \hat{\beta}]$$

Так как матрица D в методе инвариантного погружения играет роль весовой матрицы при отклонении вектора приближенного решения от оптимального, то в данном случае для переменных β_i соответствующие компоненты D характеризуют степень их отклонения от коэффициентов разложения истинной плотности (компоненты D_0, соответственно, – отклонения параметров в начальный момент). Существенным достоинством рассмотренного подхода, несмотря на формирование приближенного решения, является возможность синтеза оптимального управления $\mathbf{U}_{\text{opt}}$ в реальном времени.

4. Пример. Для иллюстрации эффективности использования предложенного подхода рассмотрим следующий пример.

Объект управления описывается уравнением

$$\dot{x} = -ax^3 + u + V, \quad x(t_0) = 0$$

где $V(t)$ – белый центрированный гауссовский шум интенсивности D_V.

Синтез управления u будем осуществлять, исходя из условия минимального отклонения координаты x от начального состояния $x(t_0)$ на интервале $T = [t_0, t_k]$ при минимуме затрат на управляющее воздействие. Традиционный подход позволяет решить данную задачу на основе среднеквадратичного критерия [2]

$$J_1 = M\left\{\int_{t_0}^{t_k} [x^2(\tau) + K^2u^2(\tau)]d\tau\right\}, \quad K = \text{const}$$

а разработанный – на основе критерия обеспечения максимума вероятности существования x в заданной окрестности $x_* = [X_{\min}, X_{\max}]$ точки x_0 (который в силу неравенства Чебышева позволяет обеспечить потенциально большую точность управления переменной x), т.е. на основе минимизации критерия

$$J_2 = \int_{x_*} \int_{t_0}^{t_k} [-\rho(x,\tau) + K^2u^2(x,\tau)]\, d\tau dx$$

В первом случае оптимальное управление определяется как [2]

$$u_{\text{opt}} = \frac{1}{2K^2} M[\lambda] = \frac{1}{2K^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \lambda\rho(x,\lambda,\tau) d\lambda dx$$

где λ – сопряженная переменная, а система канонических сопряженных уравнений, вытекающая из стохастического гамильтониана, имеет вид

$$\dot{x} = -ax^3 + u_{\text{opt}} + V_t, \quad x(t_0) = x_0$$

$$\dot{\lambda} = 3a\lambda x^2 + 2x, \quad \lambda(t_k) = 0$$

что позволяет записать уравнение для плотности $\rho = \rho(x, \lambda, t)$ следующим образом:

$$\frac{\partial\rho}{\partial t} = -\left\{\frac{\partial}{\partial x}\left[\left(\frac{1}{2K^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \lambda\rho dx d\lambda - ax^3\right)\rho\right] + \frac{\partial}{\partial\lambda}[(3a\lambda x^2 + 2x)\rho]\right\} + \frac{D_V}{2}\frac{\partial^2\rho}{\partial x^2} \tag{4.1}$$

Следовательно, точное решение поставленной задачи сводится в данном случае к решению двумерного интегродифференциального уравнения (4.1), причем с неопределенными краевыми условиями (в силу одновременной неопределенности переменных x и λ в моменты времени соответственно t_0 и t_k). Последнее обстоятельство приводит к принципиальной невозможности точного синтеза статистически оптимального управления на основе традиционного подхода (в отличие от разработанного). Классическим путем устранения данного противоречия является гауссовская аппроксимация плотности $\rho(x, \lambda, t)$, приводящая в рассматриваемом случае к сопряженной системе для оценок средних:

$$\dot{\hat{x}} = -a\hat{x}^3 + \frac{1}{2K^2}\hat{\lambda}, \quad \hat{x}(t_0) = M(x_0) = 0 \tag{4.2}$$

$$\dot{\hat{\lambda}} = 3a\hat{\lambda}x^2 + 2\hat{x}, \quad \hat{\lambda}(t_k) = 0$$

Интегрирование системы (4.2) представляет собой решение двухточечной краевой задачи, поэтому для возможности формирования управления в реальном масштабе времени используем метод инвариантного погружения, приводящий к уравнениям для приближенной оценки $\hat{x}_*$:

$$\dot{\hat{x}}_* = -a\hat{x}_*^3 + 2D\hat{x}_*, \quad \dot{D} = 4D^2 - 9aD\hat{x}_*^2 - \frac{1}{2K^2}$$

откуда определяется приближенный закон управления

$$\hat{u}_{\text{opt}} = 2\hat{x}_* D$$

Предложенный выше альтернативный подход, в свою очередь, позволяет сформировать приближенный закон управления в виде правой части уравнения относительно переменной ϑ, составляющего единую систему с уравнениями относительно вектора приближенных коэффициентов $\boldsymbol{\beta}$ разложения плотности $\rho(x, t)$ по ортонормированной системе функций φ:

$$\begin{vmatrix} \dot{\hat{\vartheta}} \\ \dot{\hat{\boldsymbol{\beta}}} \end{vmatrix} = \begin{vmatrix} 0 \\ \int\limits_{x_*} \varphi\{B_1(\varphi,x)\hat{\boldsymbol{\beta}} + 3x^2 a\varphi^T\hat{\boldsymbol{\beta}}\}dx \end{vmatrix} + D\int\limits_{x_*} \varphi(\varphi^T\hat{\boldsymbol{\beta}})^{-1}dx$$

$$\dot{D} = 2\int\limits_{x_*}\varphi\{B_1(\varphi,x) + 3x^2 a\varphi^T\}dxD + 2D\int\limits_{x_*}\varphi\{\varphi^T(B_1(\varphi,x)\hat{\boldsymbol{\beta}}(\varphi^T\hat{\boldsymbol{\beta}})^{-1} + 3ax^2)\}dx +$$

$$+\frac{2}{K^2}\int\limits_{x_*}\varphi\left[\frac{\partial}{\partial x}\left\{(\varphi^T\hat{\boldsymbol{\beta}})^2\frac{\partial}{\partial x}(\varphi^T[\varphi^T\hat{\boldsymbol{\beta}}])\right\}\right]dx - 2D\int\limits_{x_*}\varphi\varphi^T(\varphi^T\hat{\boldsymbol{\beta}})^{-2}dxD$$

$$B_1(\varphi,x) = \frac{D_V}{2}\frac{\partial^2\varphi^T}{\partial x^2} + ax^3\frac{\partial\varphi^T}{\partial x}$$

С целью сравнительной оценки точности обоих подходов было проведено численное моделирование субоптимального (полученного на основе метода инвариантного погружения) управления рассмотренным объектом при

$$a = K = 1; \quad D_V = 1{,}7; \quad x_0 = 0, \quad x_* = [-2{,}3; 2{,}3]; \quad T = [0, 100]c$$

$$\varphi = (\cos(\omega_0 x), \sin(\omega_0 x), \cos(2\omega_0 x), \sin(2\omega_0 x)), \quad \omega_0 = \pi / 2{,}3$$

реализованного для 30 стохастических траекторий движения объекта.

Интегрирование уравнений осуществлялось методом Рунге–Кутта третьего порядка с шагом 0,03 с. Оценка точности управления производилась путем усреднения по ансамблю реализаций среднемодульных отклонений отдельно взятых траекторий движения от границ интервала x_* в течение времени T. По окончании моделирования было установлено, что точность управления на основе разработанного подхода выше традиционного более чем в 2 раза, что в свою очередь позволяет сделать вывод о возможности эффективного использования предложенного метода при синтезе управления реальными стохастическими объектами.

ЛИТЕРАТУРА

1. *Афанасьев В.Н., Колмановский В.Б., Носов В.Р.* Математическая теория конструирования систем управления. М.: Высш. шк., 1989. 447 с.
2. *Казаков И.Е.* Статистическая теория систем управления в пространстве состояний. М.: Наука, 1975. 432 с.
3. *Сиразетдинов Т.К.* Оптимизация систем с распределенными параметрами. М.: Наука, 1977. 479 с.
4. *Первачев С.В., Перов А.И.* Адаптивная фильтрация сообщений. М.: Радио и связь, 1991. 160 с.

Ростов-на-Дону

Поступила в редакцию
24.II.1995