УДК 519.865

А.В.Колногоров

О МНОГОВАРИАНТОМ ЦЕЛЕСООБРАЗНОМ УПРАВЛЕНИИ СРЕДНИМ УРОВНЕМ СЛУЧАЙНЫХ ПОМЕХ

For random noise admitting the *K*-variant choice, an algorithm of rational control which constrains the limiting mean variance to its *K*-fold least value, is proposed.

Рассмотрим управляемый случайный процесс ξ_t, $t = 1,2,3,\ldots$, значения которого в каждый момент времени зависят только от текущего выбранного варианта и не зависят от остальной предыстории — откликов процесса на ранее применявшиеся варианты. Число вариантов конечно и равно K ($K \ge 2$). Будем предполагать, что рассматриваемый случайный процесс является целочисленным и равномерно ограниченным. Таким образом, если в момент времени t выбирается вариант l, то распределение случайной величины ξ_t задано совокупностью вероятностей $p_l(i) = \Pr_l\{\xi_t = i\}$, $l = 1,\ldots,K$, $i\in I$, где $I = (0,\pm1,\pm2,\ldots,\pm L)$, причем выполнены равенства

$$\sum_{i\in I} p_l(i) = 1,\ E_l\xi_t = \sum_{i\in I} i\, p_l(i) = 0,\ D_l\xi_t = \sum_{i\in I} i^2 p_l(i) = \sigma_l^2,\ \ l = 1,\ldots,K.$$

Здесь E_l, D_l обозначают математическое ожидание и дисперсию соответственно при условии выбора варианта с номером l. Рассматриваемые случайные величины интерпретируются как помехи с нулевыми математическими ожиданиями. Дисперсии помех конечны, характеризуют их амплитуды, и предполагаются априори неизвестными.

Цель состоит в снижении среднего уровня помех, т.е. их дисперсии, поэтому будем предполагать, что не все значения $\sigma_l^2,\ l = 1,\ldots,K$ одинаковы. В этом случае простая стратегия управления, предписывающая на каждом шаге равновероятный выбор вариантов, обеспечивает равенство средней дисперсии величине $D\xi_t = K^{-1}\sum_{l=1}^{K}\sigma_l^2$, которая меньше максимально возможного из значений $\sigma_l^2,\ l = 1,\ldots,K.$ Под целесообразным управлением понимается такое, при котором средняя дисперсия помех ниже, чем для указанной простой стратегии. Отметим адаптивную постановку задачи: поскольку значения дисперсий являются неизвестными, они должны так или иначе оцениваться в процессе управления. В качестве близкой постановки отметим традиционную задачу целесообразного поведения [1,2], в которой минимизируется среднее значение математического ожидания случайных величин, интерпретируемых уже как потери (если же как доходы, то целью является их максимизация).

Для случая двухвариантного управления в [3] предложен алгоритм на основе кусочно-однородного случайного блуждания на отрезке с отражающими экранами на концах, для которого доказано, что он обеспечивает целесообразное управление. Этот алгоритм не имеет очевидного обобщения на многовариантный случай, а указанный в [3] возможный способ его обобщения не является верным. Ниже предлагается алгоритм, применимый при всех значениях $K \ge 2$, который для двухвариантного случая эквивалентен рассмотренному в [3].

Опишем этот алгоритм. Пусть $R > 0$ — некоторое достаточно большое число и пусть $\eta_t^{(1)},\eta_t^{(2)},\ldots,\eta_t^{(K)}$ описывают K процессов случайного блуждания. При $t = 1$ все процессы находятся в начале координат, т.е. $\eta_t^{(l)} = 0,\ t = 1,\ l = 1,\ldots,K.$ Выбор вариантов выполняется следующим образом. Если в момент времени t процессы с номерами $l_1,\ldots,l_n$ удовлетворяют равенству

$$\eta_t^{(l_i)} = \min_{l=1,\ldots,K} \eta_t^{(l)},\ i = 1,\ldots,n,$$

то равновероятно выбирается один из вариантов $l_1,\ldots,l_n$. Если при этом был выбран вариант l_i, то в момент времени $t+1$ полагаем

$$\eta_{t+1}^{(l_i)} = \eta_t^{(l_i)} + \xi_t, \text{если } \eta_t^{(l_i)} + \xi_t \geq \min_{l \neq l_i} \eta_t^{(l)} - R,$$

$$\eta_{t+1}^{(l_i)} = \eta_t^{(l_i)}, \text{если } \eta_t^{(l_i)} + \xi_t < \min_{l \neq l_i} \eta_t^{(l)} - R.$$

Состояния остальных случайных процессов не меняются. Таким образом, в каждый момент времени активен только один из тех процессов, которые находятся левее остальных на числовой прямой. С течением времени процессы коллективно смещаются вправо, причем их максимальное удаление друг от друга не превосходит величины $R + L$: последний может отстать от предпоследнего не более чем на R, а обогнать его не более чем на L. Это обстоятельство позволяет прямую замкнуть в кольцо достаточно большой длины, указав на нем направление случайного блуждания и рассматривать процессы как конечную однородную марковскую цепь. Эта цепь непериодическая, так как она может в течение ряда тактов не менять состояние.

Обозначим через $q_l(t)$ вероятность выбора l-го варианта в момент времени t. Из свойств марковской цепи следует, что существуют предельные вероятности

$$\pi_l(R) = \lim_{t\to\infty} q_l(t), \quad l = 1,\ldots,K.$$

При этом будут выполнены равенства

$$\lim_{R\to\infty} \pi_l(R) = \sigma_l^{-2}\left(\sum_{l=1}^{K} \sigma_l^{-2}\right)^{-1}, \ l = 1,\ldots,K. \qquad (1)$$

Поэтому предельная средняя дисперсия помех

$$\sigma^2 = K\left(\sum_{l=1}^{K} \sigma_l^{-2}\right)^{-1}, \qquad (2)$$

т.е. равна среднему гармоническому дисперсий σ_l^2, $l = 1,\ldots,K.$. Если не все дисперсии равны между собой, то это означает, что выполнены неравенства

$$\sigma^2 < K^{-1}\sum_{l=1}^{K}\sigma_l^2, \ \ \sigma^2 < K \min_{l=1,\ldots,K} \sigma_l^2, \qquad (3)$$

т.е. управление является целесообразным, и предельное значение средней дисперсии не превосходит ее K-кратного минимального значения.

Замечание. При $K = 2$ разность $\eta_t^{(1)} - \eta_t^{(2)}$ описывает кусочно-однородное случайное блуждание на отрезке $[-R,R]$ с отражающими экранами на концах: при $\eta_t^{(1)} < \eta_t^{(2)}$ — случайное блуждание на полуинтервале $[-R,0)$ с дисперсией σ_1^2, а при $\eta_t^{(1)} > \eta_t^{(2)}$ — случайное блуждание на полуинтервале $(0,R]$ с дисперсией σ_2^2. Поэтому при $K = 2$ предложенный алгоритм эквивалентен рассмотренному в [3], для которого свойства (1) – (3) доказаны строго. При $K > 2$ разность $\eta_t^{(l)} - \eta_t^{(k)}$, $k,l = 1,\ldots,K$, $k \neq l$, описывает кусочно-однородное случайное блуждание на отрезке $[-R - L, R + L]$, причем отражение может происходить в точках, отстоящих от границы не более чем на $2L$. Так как величина L фиксирована, то при $R\to\infty$ отношение $\pi_l(R)/\pi_k(R)$ стремится к σ_k^2/σ_l^2, откуда следуют свойства (1) – (3).

В таблице приведены результаты численного моделирования при $K = 3$, если $p_1(1) = p_1(-1) = 1/2$, $p_2(1) = 3/4$, $p_2(-3) = 1/4$, $p_3(1) = 2/3$, $p_3(-2) = 1/3$. В этом случае $\sigma_1^2 = 1$, $\sigma_2^2 = 3$, $\sigma_3^2 = 2$. Продолжительность моделирования характеризуется величиной T. В таблице представлены величины T_1/T_2, T_1/T_3, равные отношениям времен, в течение которых применялись варианты, и характеризующие отношения вероятностей $\pi_1(R)/\pi_2(R)$, $\pi_1(R)/\pi_3(R)$.

Продолжит. моделирования	$R = 20$		$R = 50$		$R = 80$	
	T_1/T_2	T_1/T_3	T_1/T_2	T_1/T_3	T_1/T_2	T_1/T_3
$T = 2\cdot10^8$	2,90	1,87	2,96	1,95	2,99	1,98
$T = 0,5\cdot10^8$	2,91	1,88	2,96	1,94	3,00	1,96

В заключение отметим, что указанный алгоритм может быть использован в системах управления, использующих только простейшие вычислительные операции — сложение и вычитание.

1. Срагович В.Г. Адаптивное управление. М.: Наука, 1981. 319 с.
2. Назин А.В., Позняк А.С. Адаптивный выбор вариантов. М.: Наука, 1986. 288 с.
3. Колногоров А.В. // Автоматика и телемеханика. 2000. №1. С.70-80.