УДК 519.81

Профессор Ю.В. Бугаев, доцент Б.Е. Никитин,
ассистент А.С. Чайковский
(Воронеж. гос. ун-т инж. технол.) кафедра информационных технологий,
моделирования и управления, тел. (473) 255-25-50

Приближенный метод построения коллективного выбора

Рассмотрен метод нахождения оценок коэффициентов функции полезности, которая выступает в качестве структуры механизма выбора, порождающего наблюдаемую по результатам групповой экспертизы на ограниченной выборке альтернатив коллективную функцию выбора. Приведен численный пример, иллюстрирующий работу предложенного метода.

In article the method of a finding of estimations of coefficients of function of usefulness which represents itself as structure of the mechanism of the choice generating observed by results of group expertise on limited sampling of alternatives a collective choice function is considered. The numerical example illustrating operation of the offered method is resulted.

Ключевые слова: функция выбора, бинарные отношения, случайная величина.

Как правило, задача коллективного выбора формулируется так. Существует набор многокритериальных альтернатив и есть коллектив экспертов, от которых требуется упорядочить в соответствии со своими предпочтениями все альтернативы. На основании результатов опроса экспертов требуется их индивидуальные предпочтения преобразовать в некий единый коллективный выбор, т.е. провести голосование.

Традиционные процедуры голосования требуют вовлечения в процесс сравнения и оценки всего имеющегося набора альтернатив, поэтому при большой мощности исходного набора вариантов они становятся непригодны.

Принципиально иной подход содержится в методе экстраполяции экспертных оценок (МЭЭО), предложенном в [1, 2] группой ученых ВГУИТ. В этом методе на основе экспертного сравнения альтернатив из некоторой небольшой обучающей выборки строится система равенств и неравенств, описывающая область возможных значений вектора коэффициентов функции обобщенного критерия (ОК). Один из вариантов МЭЭО [3] предполагает, что в оценивании альтернатив могут участвовать несколько экспертов, расхождения в их мнении интерпретируются как случайные ошибки, а итоговые оценки коэффициентов функции ОК определяются на основе принципа максимального правдоподобия (ММП). Таким образом, данный вариант МЭЭО может быть использован как механизм голосования при коллективном выборе.

Метод экстраполяции экспертных оценок. Рассмотрим применение ММП в экстраполяции экспертных оценок. Исходной предпосылкой для МЭЭО является существование функции обобщенного критерия (ОК)

$$F(x,b)=\sum b_u f_u(x)=b^T f(x), \qquad (1)$$

где $f_u(x)$ – известные функции; b_u – неизвестные параметры.

В общем случае результат ранжирования обучающей выборки из m альтернатив каждым r-м экспертом можно представить в виде матричного неравенства

$$C^{(r)}w \geq 0, \qquad (2)$$

где w – вектор полезностей альтернатив; $C^{(r)}$, $r=1,\ldots,n$ – матрица, отображающая вариант упорядочения, выбранный r-м экспертом; n – число экспертов.

Система (2) соответствует наступлению определенного случайного события, вероятность которого определяется по формуле

$$P_r(\theta)=\int_{D_r} g(x,a,\sigma)dx, \qquad D_r=\left\{x\in E^m \,\middle|\, C^{(r)}x\geq 0\right\}, \qquad (3)$$

где $g(x,a,\sigma)$ – плотность m-мерного нормального распределения вида

$$g(x,a,\sigma)= =\frac{1}{(2\pi)^{m/2}\sigma^m}\exp\left[-\frac{1}{2\sigma^2}(x-a)^T(x-a)\right].$$

Функция правдоподобия в соответствии с формулой полиноминального распределения примет вид

$$L(k_1,\cdots,k_s,\theta)=\frac{n!}{k_1!\ldots k_s!}P_1^{k_1}(\theta)\ldots P_s^{k_s}(\theta), \quad (4)$$

где k_r – количество экспертов, выбравших r-й вариант упорядочения; $\theta=(a_1,\ldots,a_m,\sigma)^T$ – вектор оцениваемых параметров. В результате максимизации выражения (4) получаются точечные оценки коэффициентов ОК.

Обоснование данного метода и описание вычислительных процедур для простейших вариантов групповой экспертизы было дано одним из авторов настоящей статьи в [4]. При попытках обобщения метода на случай произвольной стратегии экспертного сравнения альтернатив выборки возникают проблемы вычислительного характера, т.к. при определении значений требуемых вероятностей приходится вычислять несобственные интегралы высокой кратности по области сложной структуры. Поэтому возникла необходимость в выработке приближенных алгоритмов ее решения.

В [5] был предложен подход, включающий в себя комбинацию эвристической процедуры поиска начального приближения и метода Монте-Карло для вычисления несобственных интегралов, зависящих от оптимизируемых величин как от параметров.

Авторами данной статьи был разработан соответствующий программный продукт. Вычислительные эксперименты показали, что применение обычного метода Монте-Карло в связи с наличием случайной составляющей в процедуре вычисления интегралов неизбежно приводит к тому, что значение функции вычисляется со случайной погрешностью, что, в свою очередь, негативно отражается на точности решения и скорости сходимости.

Одним из способов преодоления данной проблемы является использование метода Соболя [6]. Проведенные эксперименты показали, что данный способ вычисления интеграла обеспечивает лучший порядок сходимости и позволяет получить более гладкую поверхность.

Таким образом, ни один из перечисленных выше методов не позволяет избежать негладкости целевой функции.

Предлагается способ преодоления данной проблемы путем применения метода Ньютона. Однако чтобы им воспользоваться необходимо вычислить первую и вторую производные. Рассмотрим решение в аналитическом виде.

Метод Ньютона. В общем случае вероятность ранжирования вычисляется по формуле

$$P(b)=\lambda\int_{Cx\geq 0}\exp\left[-\frac{1}{2}(x-Zb)^T(x-Zb)\right]dx,$$

где λ – константа, определяемая условием нормировки плотности распределения; Z – матрица координат альтернатив в пространстве базисных функций f_j, т.е. $Z_{ij}=f_j(x^i)$. Вычислим векторную логарифмическую производную $\frac{d\ln(P)}{db}$. Имеем: $\frac{d\ln(P)}{db}=\frac{1}{P}\nabla_b P$.

В свою очередь,

$$\nabla_b P=\lambda\int_{Cx\geq 0}\nabla_b\exp\left[-\frac{1}{2}(x-Zb)^T(x-Zb)\right]dx=$$

$$=\lambda\int_{Cx\geq 0}\exp\left[-\frac{1}{2}(x-Zb)^T(x-Zb)\right]\cdot$$

$$\nabla_b\left[-\frac{1}{2}(x-Zb)^T(x-Zb)\right]dx.$$

Показатель экспоненты можно преобразовать к виду

$$-\frac{1}{2}(x-Zb)^T(x-Zb)=$$

$$=-\frac{1}{2}\left(x^Tx-x^TZb-b^TZ^Tx+b^TZ^TZb\right)=$$

$$=\frac{1}{2}\left(-x^Tx+\left(Z^Tx\right)^Tb+b^TZ^Tx-b^TZ^TZb\right).$$

Известны формулы матричного дифференцирования:

$$\nabla_y(a^Ty)=\nabla_y(y^Ta)=a,\ \nabla_y(y^TAy)=2Ay,$$

где a – постоянный вектор; A – квадратная симметричная матрица. Отсюда

$$\nabla_b\left(-\frac{1}{2}\left[-x^Tx+\left(Z^Tx\right)^Tb+b^TZ^Tx-b^TZ^TZb\right]\right)=$$

$$=Z^Tx-Z^TZb=Z^T(x-Zb).$$

Окончательно $\frac{d\ln(P)}{db}=\frac{\lambda}{P}Z^T$,

$$\int_{Cx\geq 0}\exp\left[-\frac{1}{2}(x-Zb)^T(x-Zb)\right]\cdot(x-Zb)dx. \quad (5)$$

Выясним теоретико-вероятностный смысл формулы (5). Рассмотрим функцию

$$f(x)=\frac{1}{P}\lambda\exp\left[-\frac{1}{2}(x-Zb)^T(x-Zb)\right]=$$

$$=\frac{\exp\left[-\frac{1}{2}(x-Zb)^T(x-Zb)\right]}{\int\limits_{Cx\geq 0}\exp\left[-\frac{1}{2}(x-Zb)^T(x-Zb)\right]dx}.$$

Её можно интерпретировать как плотность усечённого многомерного нормального распределения с параметрами (Zb, I), определённого в области $Cw \geq 0$. Тогда интеграл

$$\frac{\lambda}{\mathrm{P}}\int\limits_{Cx\geq 0}\exp\left[-\frac{1}{2}(x-Zb)^T(x-Zb)\right]\cdot(x-Zb)dx$$

представляет собой среднее значение на области $C(v+Zb)\geq 0$ m-мерной усечённой нормальной случайной величины v параметрами $(0, I)$, определённой в этой области. Обозначим $\bar{v}$ – вектор этого среднего значения. Тогда интеграл (5) – не что иное, как вектор $Z^T\bar{v}$. Его приближённое значение легко вычислить с помощью имитации по алгоритму.

1. Получить N независимых m-мерных векторов v^i, распределённых по закону $N(0, I)$.

2. Отобрать из них множество D векторов, удовлетворяющих условию $C(v^i + Zb)\geq 0$.

3. Вычислить вектор $\bar{v}$ среднего значения векторов v^i, попавших в D.

4. Вычислить $Z^T\bar{v}$. Это и будет приближённым значением $\frac{d\ln(P)}{db}$.

Этот алгоритм можно использовать при максимизации функции правдоподобия, если интегралы (2) вычислять методом Монте-Карло. Как было сказано ранее, при вычислении интегралов (2) методом Монте-Карло невозможно избежать негладкости целевой функции. Поэтому численное вычисление её градиента в стандартных процедурах оптимизации связано с большими погрешностями. Использование же алгоритма позволяет значительно уменьшить эти погрешности. Тогда для оптимизации можно воспользоваться программой, которая использует пользовательские формулы вычисления градиента целевой функции. Аналогично можно вычислять и матрицу вторых частных производных и использовать её при оптимизации.

Рассмотрим нахождение матрицы вторых производных $\frac{d^2\ln(P)}{db^2}$ для частного случая. Пусть имеем функцию вида

$$f(b_1,b_2)=\exp\left[-\frac{1}{2}(x-Zb)^T(x-Zb)\right],$$

где $Z=\begin{bmatrix} z_{11} & z_{12}\\ z_{21} & z_{22}\end{bmatrix}$; $x=\begin{bmatrix} x_1\\ x_2\end{bmatrix}$; $b=\begin{bmatrix} b_1\\ b_2\end{bmatrix}$.

Тогда

$$\frac{d^2\ln(P)}{db^2}=\frac{d}{db}\left(\frac{1}{P}\nabla_b P\right), \qquad (6)$$

где $\nabla_b P = f(b_1,b_2)Z^T(x-Zb)$. Последнее выражение представим в виде

$$f(b_1,b_2)Z^T(x-Zb)=$$
$$=f(b_1,b_2)\begin{bmatrix} z_{11} & z_{21}\\ z_{12} & z_{22}\end{bmatrix}\begin{bmatrix} x_1-z_{11}b_1-z_{12}b_2\\ x_2-z_{21}b_1-z_{22}b_2\end{bmatrix}=$$
$$=f(b_1,b_2)\begin{bmatrix} z_{11}(x_1-z_{11}b_1-z_{12}b_2)+z_{21}(x_2-z_{21}b_1-z_{22}b_2)\\ z_{12}(x_1-z_{11}b_1-z_{12}b_2)+z_{22}(x_2-z_{21}b_1-z_{22}b_2)\end{bmatrix}$$

Для компактности дальнейших записей введем следующее обозначение:

$$V=\begin{bmatrix} z_{11}(x_1-z_{11}b_1-z_{12}b_2)+z_{21}(x_2-z_{21}b_1-z_{22}b_2)\\ z_{12}(x_1-z_{11}b_1-z_{12}b_2)+z_{22}(x_2-z_{21}b_1-z_{22}b_2)\end{bmatrix}$$

Таким образом, выражением (6) имеет вид

$$\frac{d}{db}\left(\frac{1}{P}f(b_1,b_2)V\right)=$$
$$=\left\{\left(-\frac{1}{P^2}\frac{\partial P}{\partial b_1}f(b_1,b_2)V+\frac{1}{P}\frac{\partial f(b_1,b_2)}{\partial b_1}V+\right.\right.$$
$$\left.+\frac{1}{P}f(b_1,b_2)\begin{bmatrix} -z_{11}z_{11} & -z_{21}z_{21}\\ -z_{12}z_{11} & -z_{21}z_{22}\end{bmatrix}\right),$$
$$\left(-\frac{1}{P^2}\frac{\partial P}{\partial b_2}f(b_1,b_2)V+\frac{1}{P}\frac{\partial f(b_1,b_2)}{\partial b_2}V+\right.$$
$$\left.\left.+\frac{1}{P}f(b_1,b_2)\begin{bmatrix} -z_{11}z_{11} & -z_{21}z_{21}\\ -z_{12}z_{11} & -z_{21}z_{22}\end{bmatrix}\right)\right\}=$$
$$=\frac{1}{P}\left\{-\frac{1}{P}\left(\frac{\partial P}{\partial b_1}f(b_1,b_2)V,\frac{\partial P}{\partial b_2}f(b_1,b_2)V\right)+\right.$$
$$\left.+\left(\frac{\partial f(b_1,b_2)}{\partial b_1}V,\frac{\partial f(b_1,b_2)}{\partial b_2}V\right)-f(b_1,b_2)Z^TZ\right\}.$$

Окончательно получим

$$\frac{d^2\ln(P)}{db^2}=\frac{1}{P}\left\{-\frac{1}{P}\left((\nabla_b P)(\nabla_b P)^T\right)+\right.$$
$$\left.+Z^T(x-Zb)(\nabla f(b_1,b_2))^T-f(b_1,b_2)Z^TZ\right\}.$$

Так, матрица вторых производных имеет следующий вид:

$$\frac{d^2\ln(P)}{db^2}=\frac{d}{db}\left(\frac{1}{P}\nabla_b P\right)=\frac{1}{P}\left(-\frac{1}{P}(\nabla_b P)(\nabla_b P)^T+\right.$$
$$\left.+Z^T(x-Zb)(\nabla_b f(b))^T-f(b)Z^TZ\right),$$

где $f(b) = \exp\left(-\frac{1}{2}(x - Zb)^T (x - Zb)\right)$.

Покажем применение данного метода на следующем примере. Пусть дана ступенчатая функция

$$F(a) = 2\ln\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} e^{-\frac{1}{2}t^2}\right) + \ln\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-a} e^{-\frac{1}{2}t^2}\right),$$

необходимо найти ее максимум. В результате аппроксимации методом Соболя (127 точек) была получена функция представленная на рис. 1. Данная функция является негладкой, потому непосредственный поиск максимума не дает верного решения. Если же воспользоваться методом Ньютона с начальным приближением 1,9, то уже на 4-й итерации получим решение с точностью порядка 10^{-3}. На рис. 2 изображен результат 1-й итерации, на рис. 3 – результат 2-й итерации.

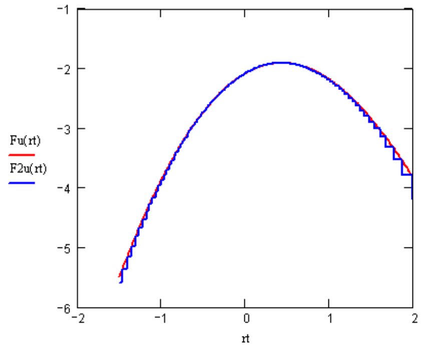

Рис 1

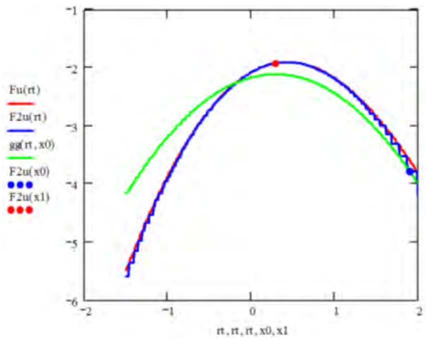

Рис. 2

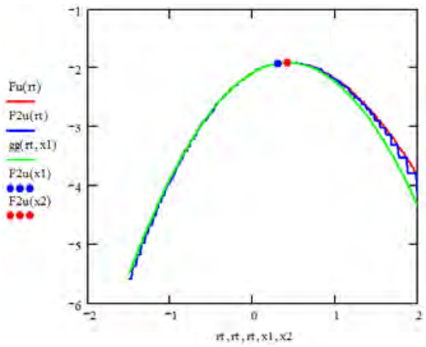

Рис. 3

Таким образом, предложенный метод позволяет получить более точное решение, поскольку в сравнении с другими методами дает возможность построить достаточно гладкую целевую функцию.

Работа выполнена в рамках ГК от 21 апреля 2011 г. № 16.516.11.6040 «Проведение проблемно-ориентированных поисковых исследований в области создания систем мониторинга и управления энергопотреблением в зданиях и сооружениях».

ЛИТЕРАТУРА

1. Пустыльник, Е.И. Использование линейной модели для экстраполяции экспертных оценок [Текст] / Е.И. Пустыльник, В.В. Сысоев, М.С. Чирко // Автоматизация проектирования: сб. науч. тр.– М., 1981. – С. 46-50.

2. Пустыльник, Е.И. Об одном методе экстраполяции экспертных оценок [Текст] / Е.И. Пустыльник, В.В. Сысоев, М.С. Чирко // Экономика и мат. методы. – 1983. –Вып. 4. – С. 716-717.

3. Десятов, Д.Б. Принятие решений на основе экспертных оценок с использованием метода максимального правдоподобия [Текст] / Д.Б. Десятов, В.В. Сысоев, М.С. Чирко // Автоматизация проектирования производственных систем: сб. науч. тр. - Воронеж, 1984. – С. 32-36.

4. Бугаев, Ю.В. Экстраполяция экспертных оценок в оптимизации технологических систем [Текст] / Ю.В. Бугаев // Изв. АН. Теория и системы управления. – 2003. – № 3. – С. 90–96.

5. Бугаев, Ю.В. Приближенный метод синтеза моделей выбора на основе экстраполяции экспертных оценок [Текст] / Ю.В. Бугаев, И.Е. Медведкова, Б.Е. Никитин, А.С. Чайковский // Вестник ТГТУ. -2009. -Том 15. -№ 4. -С. 766-776.

6. Соболь, И.М. Выбор оптимальных параметров в задачах со многими критериями [Текст] / учеб. пособие для вузов / И.М. Соболь, Р.Б. Статников. – 2-е изд., перераб. и доп. – М.: Дрофа, 2006. – 175 с.