МЕТОД ФУНКЦИЙ ОБРАТНЫХ СВЯЗЕЙ

Метод функций обратных связей

Рассмотрим теперь другой способ сведения исходной задачи математического программирования к решению системы нелинейных уравнений с инструментальным параметром.

Пусть достаточно гладкая штрафная функция $P(\tau, s)$ кроме условия

$$\lim_{\tau \to +0} P(\tau, s) = \begin{cases} 0\,, & s < 0, \\ +\infty\,, & s > 0, \end{cases}$$

удовлетворяет также и неравенствам

$$\frac{\partial P}{\partial s} > 0 \quad \text{и} \quad \frac{\partial^2 P}{\partial s^2} > 0 \qquad \forall s \in \Re,$$

Заметим, что, например, стандартная квадратичная штрафная функция вида $P(\tau, s) = \begin{cases} 0, & s \le 0, \\ \dfrac{s^2}{2\tau}, & s > 0 \end{cases}$ не удовлетворяет этим условиям.

Если $A(\tau, x)$ достаточно гладкая в E^n функция, имеющая при фиксированном $\tau > 0$ изолированную точку локального максимума $\bar{x}(\tau)$, то неявно заданная *функция* $\bar{x}(\tau)$ определяется равенством:

$$\frac{\partial A}{\partial \xi_k} = 0 \qquad \forall k = [1, n],$$

Из предположений о свойствах (в данном случае, непрерывной дифференцируемости) штрафной функции следует существование пределов

$$\lim_{\tau \to +0} \frac{\partial P}{\partial s}(\tau, f_i(\bar{x}(\tau))) \qquad \forall i = [1, m].$$

Тогда из сравнения условия стационарности $A(\tau, x)$, записанное в форме

$$\operatorname{grad} F(x) - \sum_{i=1}^{m} \frac{\partial P}{\partial s}(\tau, f_i(\bar{x}(\tau))) \cdot \operatorname{grad} f_i(x) = o$$

с утверждением теоремы Каруша–Куна–Таккера, о том, что существуют и единственны неотрицательные числа − множители Лагранжа $\lambda_i^* \geq 0 \quad \forall i = [1, m]$ таких, что

$$\operatorname{grad} F(x) - \sum_{i=1}^{m} \lambda_i^* \operatorname{grad} f_i(x) =, o$$

можно прийти к заключению, что для изолированного локального решения исходной задачи математического программирования x^* справедливы следующие предельные соотношения:

$$\lim_{\tau \to +0} \frac{\partial P}{\partial s}(\tau, f_i(\bar{x}(\tau))) = \lambda_i^* \quad \forall i = [1, m]$$

Следовательно, величины $\frac{\partial P}{\partial s}(\tau, f_i(\bar{x}(\tau)))$ можно использовать как приближенные оценки значений множителей Лагранжа.

Воспользуемся полученной оценкой для решения методом гладких штрафных функций следующей пары взаимодвойственных задач.

Пусть функции $F(x),\ f_i(x) \quad \forall i=\overline{1,m}$ *линейны* в E_+^n − неотрицательном ортанте E^n.

Тогда исходная (прямая) задача имеет вид

$$F(x)=\sum_{j=1}^{n}\sigma_j\xi_j \to \max \quad x\in E_+^n$$

$$\textit{при условиях} \quad f_i(x)=-\beta_i+\sum_{j=1}^{n}\alpha_{ij}\xi_j \le 0 \quad \forall i=\overline{1,m},$$

где $\sigma_j,\beta_i,\alpha_{ij}\ \forall j=\overline{1,n}, \forall i=\overline{1,m}$ − константы. Обозначим ее решение как $x^*=(\xi_1^*,\xi_2^*,\ldots,\xi_m^*)^{\mathrm{T}}$.

Двойственная к исходной линейная задача в этом случае будет

$$G(\lambda)=\sum_{i=1}^{m}\beta_i\lambda_i \to \min \quad \lambda\in E_+^m$$

$$\textit{при условиях} \quad g_j(\lambda)=-\sigma_j+\sum_{i=1}^{m}\alpha_{ij}\lambda_i \ge 0 \quad \forall j=\overline{1,n},$$

где $\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_m)^{\mathrm{T}}$. Решение двойственной задачи обозначим как $\lambda^*=(\lambda_1^*,\lambda_2^*,\ldots,\lambda_m^*)^{\mathrm{T}}$.

Пусть вспомогательная функция прямой задачи

$$A_P(\tau,x)=F(x)-\sum_{i=1}^{m}P\big(\tau,f_i(x)\big)-\sum_{j=1}^{n}P\big(\tau,-x_j\big)\,,$$

Для двойственной задачи вспомогательная функция (подлежащая минимизации) будет иметь вид

$$A_D(\tau,\lambda)=G(\lambda)+\sum_{j=1}^{n}P\big(\tau,-g_j(\lambda)\big)+\sum_{i=1}^{m}P\big(\tau,-\lambda_i\big)$$

Обозначим их экстремальные точки как:

$$\tilde{x}(\tau)=\arg\max_{x}A_P(\tau,x) \quad \text{и} \quad \hat{\lambda}(\tau)=\arg\min_{\lambda}A_D(\tau,\lambda)\,.$$

Предположим далее, что справедливы соотношения:

$$\xi_j^* = \lim_{\tau \to +0} \tilde{\xi}_j(\tau) \quad \forall j = \overline{1,n} \quad \text{и} \quad \lambda_i^* = \lim_{\tau \to +0} \hat{\lambda}_i(\tau) \quad \forall i = \overline{1,m},$$

так же как и

$$\lambda_i^* = \lim_{\tau \to +0} \frac{\partial P}{\partial s}\left(\tau, f_i\left(\tilde{x}(\tau)\right)\right) \quad \forall i = \overline{1,m} \quad \text{и}$$

$$x_j^* = \lim_{\tau \to +0} \frac{\partial P}{\partial s}\left(\tau, -g_j\left(\hat{\lambda}(\tau)\right)\right) \quad \forall j = \overline{1,n}.$$

Для пары взаимодвойственных задач компоненты вектора – λ^* это множители Лагранжа прямой задачи, а компоненты вектора x^* – множители Лагранжа задачи двойственной.

Однако, для фиксированного $\tau > 0$, вообще говоря,

$$\frac{\partial P}{\partial s}\left(\tau, f_i\left(\tilde{x}(\tau)\right)\right) \ne \hat{\lambda}_i(\tau) \quad \forall i = \overline{1,m} \qquad \text{и}$$

$$\frac{\partial P}{\partial s}\left(\tau, -g_j\left(\hat{\lambda}(\tau)\right)\right) \ne \tilde{x}_j(\tau) \quad \forall j = \overline{1,n}.$$

Эти соотношения превращаются в верные равенства только при предельном переходе $\tau \to +0$.

Забудем теперь про метод штрафных функций и формально $\forall \tau > 0$ построим систему уравнений с неизвестными $\bar{x}(\tau)$ и $\bar{\lambda}(\tau)$ вида

$$\begin{cases} \bar{\lambda}_i = \dfrac{\partial P}{\partial s}\left(\tau, \quad f_i\left(\bar{x}(\tau)\right)\right) & \forall i = \overline{1, m}, \\ \bar{\xi}_j = \dfrac{\partial P}{\partial s}\left(\tau, -g_j\left(\bar{\lambda}(\tau)\right)\right) & \forall j = \overline{1, n}. \end{cases}$$

Можно показать, что эта система имеет решения для *любой* пары взаимодвойственных задач: совместных и несовместных, с единственным решением и с неединственным.

Заметим также, что в этой системе нет явных условий неотрицательности компонент $\bar{x}(\tau)$ и $\bar{\lambda}(\tau)$. Они здесь излишни, но почему?

Воспользуемся теперь тем, что непрерывная и строго монотонно возрастающая для $s \in (-\infty, +\infty)$ функция $\frac{\partial P}{\partial s}(\tau, s)$ имеет *обратную функцию* $Q(\tau, s)$. Эта функция $Q(\tau, s)$ также непрерывна и строго монотонно возрастает при $s \in (0, +\infty)$.

Тогда построенную систему можно записать в виде

$$\begin{cases} -\beta_i + \sum\limits_{j=1}^{n} \alpha_{ij} \overline{\xi}_j = \quad Q\left(\tau, \overline{\lambda}_i(\tau)\right) & \forall i = \overline{1, m}, \\ -\sigma_j + \sum\limits_{i=1}^{m} \alpha_{ij} \overline{\lambda}_i = -Q\left(\tau, \overline{\xi}_j(\tau)\right) & \forall j = \overline{1, n}. \end{cases}$$

Заметим, что вид данной системы оправдывает использование термина *функция обратной связи* для функции $Q(\tau, s)$.

Примерами функций обратной связи могут служить:

1) $Q(\tau, s) = \tau \ln s$ для $\frac{\partial P}{\partial s} = e^{\frac{s}{\tau}}$,

2) $Q(\tau, s) = \frac{\tau}{2}\left(s - \frac{1}{s}\right)$ для $\frac{\partial P}{\partial s} = \sqrt{\left(\frac{s}{\tau}\right)^2 + 1} + \frac{s}{\tau}$.

Для перехода к нелинейному случаю введем новую вспомогательную функцию вида

$$U(\tau, x, \lambda) = L(x, \lambda) - \sum_{j=1}^{n} R\left(\tau, \xi_j\right) + \sum_{i=1}^{m} R\left(\tau, \lambda_i\right), \qquad (*)$$

где $L(x,\lambda) = \sum_{j=1}^{n} \sigma_j \xi_j - \sum_{i=1}^{m} \lambda_i \left(-\beta_i + \sum_{j=1}^{n} \alpha_{ij} \xi_j \right)$ – регулярная функция Лагранжа прямой задачи, а $R(\tau, s) = \int\limits_{\omega(\tau)}^{s} Q(\tau, u)\, du$.

Функция $\omega(\tau)$ определена как корень уравнения $Q(\tau, \omega(\tau)) = 0$, которое имеет решение и притом единственное при каждом фиксированном. $\tau > 0$.

В этом случае векторы $\overline{x}(\tau)$ и $\overline{\lambda}(\tau)$ являются *стационарными точками* функции $U(\tau, x, \lambda)$ по компонентам векторов x и λ. Наконец, в терминах L и Q последняя система может быть записана как

$$\begin{cases} \dfrac{\partial L}{\partial \xi_j} = Q\left(\tau, \overline{\xi}_j(\tau)\right) & \forall j = \overline{1,n}, \\ \dfrac{\partial L}{\partial \lambda_i} = -Q\left(\tau, \overline{\lambda}_i(\tau)\right) & \forall i = \overline{1,m}. \end{cases}$$

Функция Лагранжа для исходной нелинейной задачи записывается в виде

$$L(x,\lambda) = F(x) - \sum_{i=1}^{m} \lambda_i f_i(x) .$$

Этот формат не зависит от того, являются ли функции $F(x),\ f_i(x) \quad \forall i = \overline{1,m}$ линейными или нет.

Поэтому естественно попытаться взять (*) в качестве определения вспомогательной функции $U(\tau, x, \lambda)$ в нелинейном случае. И рассматривать ее стационарные точки, то есть, решения системы

$$\begin{cases} \dfrac{\partial L}{\partial \xi_j} = \ \ Q\left(\tau, \overline{\xi}_j(\tau)\right) \qquad \forall j = \overline{1,n}, \\ \dfrac{\partial L}{\partial \lambda_i} = -Q\left(\tau, \overline{\lambda}_i(\tau)\right) \qquad \forall i = \overline{1,m}, \end{cases} \tag{!}$$

как приближенных решений исходной нелинейной задачи, для которых

$$\lim_{\tau \to +0} \overline{\xi}_j(\tau) = \xi_j^* \quad \forall j = \overline{1,n}, \qquad \lim_{\tau \to +0} \overline{\lambda}_i(\tau) = \lambda_i^* \quad \forall i = \overline{1,m}. \ \ .$$

Построение и последующее решение системы (!) и является *методом функций обратных связей*