ФУНКЦИИ ОБРАТНЫХ СВЯЗЕЙ

Рассмотрим вариант метода функций обратной связи, (т.е. функций реализующих в условиях оптимальности дополнительные соотношения между прямыми и двойственными переменными) для решения задач нелинейного программирования вида:

$$\begin{array}{l} F(x) \to \max, \\ \text{при условиях } \ f_i(x) \le 0 \quad \forall i = \overline{1,m}\,, \quad x_j \ge 0 \ \ \forall j = \overline{1,n}\,, \end{array} \tag{1.1}$$

где

— $x = (\, x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in E^n$ — n—мерному евклидову пространству E^n ;
— функции $F(x)$, $f_i(x)$ $\forall i = \overline{1,m}$ определены и дважды непрерывно дифференцируемы в E^n_+.

Мы будем также предполагать, что задача (1.1) имеет локальное решение x^* (быть может, не единственное) и обозначать $F(x^*)$ как F^*.

Метод функций обратных связей для нелинейных задач математического программирования

Поскольку схема решения нелинейной задачи (1.1) основана на обобщении ее реализации для задач линейного программирования, приведем вначале краткое описание линейного случая.

Пусть функции $F(x)$, $f_i(x)$ $\forall i = \overline{1,n}$ линейны, E^n_+ и E^m_+ суть неотрицательные ортанты евклидовых пространств E^n и E^m. Тогда задача (1.1) имеет вид

$$F(x) = \sum_{j=1}^{n} \sigma_j x_j \to \max \quad x \in E^n_+ \qquad \text{при условиях} \qquad f_i(x) = -\beta_i + \sum_{j=1}^{n} \alpha_{ij} x_j \le 0 \quad \forall i = \overline{1,m}. \tag{2.1}$$

Двойственная задача в данном случае имеет симметричную формулировку

$$G(\lambda) = \sum_{i=1}^{m} \beta_i \lambda_i \to \min \qquad \text{при условиях} \qquad g_j(\lambda) = -\sigma_j + \sum_{i=1}^{m} \alpha_{ij} \lambda_i \ge 0 \quad \forall j = \overline{1,n}, \tag{2.2}$$

где $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)^{\mathrm{T}} \in E^m_+$.

Применим для решения линейного варианта задачи (1.1) метод *гладких* штрафных функций, который заключается в последовательной безусловной минимизации вспомогательной функции

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

где функция $P(\tau,s)$ определяет величину «штрафа» за нарушение ограничения $s\le 0$ и удовлетворяет условиям:

1°. $\forall\tau>0$ и $\forall s$: $\quad \lim\limits_{\tau\to 0+}P(\tau,s)=\begin{cases}+\infty, & s>0\,,\\ 0, & s\le 0\,.\end{cases}$

2°. Функция $P(\tau,s)$ имеет непрерывные частные производные по всем своим аргументам до второго порядка включительно.

3°. Для всех $\tau>0$ и $\forall s$ выполнены[1] неравенства $\quad \dfrac{\partial P}{\partial s}>0\,;\quad \dfrac{\partial^2 P}{\partial s^2}>0\,.$

Для двойственной задачи вспомогательная функция метода штрафных функций будет иметь вид

$$A_{\mathrm{D}}(\tau,\lambda)=G(\lambda)+\sum_{j=1}^{n}P\Big(\tau,-g_j(\lambda)\Big)+\sum_{i=1}^{m}P\Big(\tau,(-\lambda_i)\Big).$$

[1]Заметим, что, например, стандартная квадратичная штрафная функция вида $P(\tau,s)=\begin{cases}\dfrac{s^2}{2\tau} & \text{при}\quad s\ge 0,\\ 0 & \text{при}\quad s<0\end{cases}$ не удовлетворяет этим условиям.

Пусть $\{\,\tilde{x}(\tau)\tilde{\lambda}(\tau)\,\}$ и $\{\,\breve{x}(\tau),\breve{\lambda}(\tau)\,\}$ суть соответственно стационарные точки функций $A_{\mathrm{P}}(\tau,x)$ и $A_{\mathrm{D}}(\tau,\lambda)$, определяемые при фиксированном $\tau>0$ уравнениями

$$\operatorname*{grad}_{x} A_{\mathrm{P}}\bigl(\tau,\tilde{x}(\tau)\bigr)=o \qquad \text{и} \qquad \operatorname*{grad}_{\lambda} A_{\mathrm{D}}\bigl(\tau,\breve{\lambda}(\tau)\bigr)=o, \tag{2.3}$$

Предположим далее (теория метода штрафных функций этот вариант допускает), что выполняются соотношения

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

$$\lambda_i^*=\lim_{\tau\to+0}\frac{\partial P}{\partial s}\biggl(\tau,f_i\Bigl(\tilde{x}(\tau)\Bigr)\biggr)\quad \forall i=\overline{1,m}\,, \qquad x_j^*=\lim_{\tau\to+0}\frac{\partial P}{\partial s}\biggl(\tau,-g_j\Bigl(\breve{\lambda}(\tau)\Bigr)\biggr)\quad \forall j=\overline{1,n}\,.$$

Т.е. для линейных взаимодвойственных задач (2.1) — (2.2) компоненты вектора x^* являются множителями Лагранжа для задачи (2.2), а компоненты вектора λ^* суть множители Лагранжа в задаче (2.1). При фиксированном $\tau>0$, вообще говоря,

$$\frac{\partial P}{\partial s}\biggl(\tau,f_i\Bigl(\tilde{x}(\tau)\Bigr)\biggr)\neq\breve{\lambda}_i(\tau))\quad \forall i=\overline{1,m} \qquad \text{и} \qquad \frac{\partial P}{\partial s}\biggl(\tau,-g_j\Bigl(\breve{\lambda}(\tau)\Bigr)\biggr)\neq\tilde{x}_j(\tau))\quad \forall j=\overline{1,n}\,.$$

Эти соотношения будут верными равенствами лишь в пределе при $\tau\to 0+$.

Однако данные соотношения можно объединить в одну систему уравнений. Можно показать, что для линейных задач (2.1) — (2.2) существуют вектор-функции $\overline{x}(\tau)$ и $\overline{\lambda}(\tau)$, являющиеся решениями системы

$$\begin{cases} \overline{\lambda}_i(\tau) & = & \dfrac{\partial P}{\partial s}\Big(\tau, \quad f_i\Big(\overline{x}(\tau)\Big)\Big) & \qquad \forall i = \overline{1, m}\ , \\ \overline{x}_j(\tau) & = & \dfrac{\partial P}{\partial s}\Big(\tau, -g_j\Big(\overline{\lambda}(\tau)\Big)\Big) & \qquad \forall j = \overline{1, n}\ , \end{cases} \tag{2.4}$$

для которых (в случае совместности задач (2.1) – (2.2)) верно соотношение: $\lim\limits_{\tau\to+0} F\big(\overline{x}(\tau)\big) = F^*$, а при единственности x^* и λ^* кроме того и равенства:

$$\lambda_i^* = \lim_{\tau\to+0} \frac{\partial P}{\partial s}\Big(\tau, f_i(\overline{x}(\tau))\Big) \quad \forall i = \overline{1, m}\,, \qquad\qquad x_j^* = \lim_{\tau\to+0} \frac{\partial P}{\partial s}\Big(\tau, -g_j(\overline{\lambda}(\tau))\Big) \quad \forall j = \overline{1, n}\,.$$

Тогда вектор-функции $\overline{x}(\tau)$ и $\overline{\lambda}(\tau)$, (наряду с $\tilde{x}(\tau)$, $\breve{x}(\tau)$, $\tilde{\lambda}(\tau)$ и $\breve{\lambda}(\tau)$) могут быть использованы в качестве оценочных аппроксимаций решений задач (2.1) и (2.2).

Отметим, что в системе (2.4) нет в явном виде условий, гарантирующих неотрицательность компонент векторов $\overline{x}(\tau)$ и $\overline{\lambda}(\tau)$, поскольку $\dfrac{\partial P}{\partial s} > 0$.

Теперь воспользуемся тем, что в сделанных предположениях функция $\dfrac{\partial P}{\partial s}\big(\tau, s\big)$, как непрерывно дифференцируемая и строго монотонно возрастающая по аргументу s $\forall s \in R$, имеет обратную функцию, которая также непрерывно дифференцируемая и строго монотонно возрастающая на $(0, +\infty)$. Тогда система (2.4) может быть записана как

$$\left\{ \begin{array}{rcl} f_i\big(\tau, \overline{x}(\tau)\big) & = & Q(\tau, \overline{\lambda}_i(\tau)) \quad \forall i = [1, m] \ , \\ \\ -g_j\big(\tau, \overline{\lambda}(\tau)\big) & = & Q(\tau, \overline{x}_j(\tau)) \quad \forall j = \overline{1, n} \end{array} \right. \tag{2.5}$$

где функция $Q(\tau, s) = \operatorname{inv}\left(\dfrac{\partial P}{\partial s}\big(\tau, s\big)\right)$ есть обратная для функции $\dfrac{\partial P}{\partial s}\big(\tau, s\big)$.

Построенная таким способом функция $Q(\tau, s)$ реализует *обратную связь* между прямыми и двойственными переменными в условиях оптимальности для задач (2.1) и (2.2).

Переход к нелинейному случаю выполним, введя в рассмотрение вспомогательную функцию вида

$$U(\tau,x,\lambda)=\sum_{j=1}^{n}\Big(\sigma_j x_j - R(\tau,x_j)\Big)+\sum_{i=1}^{m}\Big(\beta_i\lambda_i + R(\tau,\lambda_i)\Big)-\sum_{j=1}^{n}\sum_{i=1}^{m}\alpha_{ij}x_j\lambda_i\,, \quad \text{где} \quad R(\tau,s)=\int\limits_{\alpha(\tau)}^{s}Q(\tau,u)\,du\,, \tag{2.6}$$

а значение $\alpha(\tau)$ находится из уравнения $Q\big(\tau,\alpha(\tau)\big)=0$, которое при каждом фиксированном $\tau>0$ имеет (и притом единственное) решение, поскольку функция $Q(\tau,s)$ строго монотонно возрастающая по s и не ограниченная как снизу, так и сверху для $s\in(0,+\infty)$.

В этом случае решения системы (2.5), т.е. векторы $\overline{x}(\tau)$ и $\overline{\lambda}(\tau)$ — суть *стационарные точки* функции $U(\tau,x,\lambda)$ по совокупности $\{\,x;\,\lambda\,\}$, а сама функция (2.6) и условия ее стационарности могут быть записаны соответственно как

$$U(\tau,x,\lambda)=L(x,\lambda)-\sum_{j=1}^{n}R\big(\tau,x_j\big)+\sum_{i=1}^{m}R\big(\tau,\lambda_i\big)\,, \tag{2.7}$$

$$\begin{cases} \dfrac{\partial L}{\partial x_j}= Q(\tau,x_j) & \forall j=\overline{1,n}\,,\\ \dfrac{\partial L}{\partial \lambda_i}=-Q(\tau,\lambda_i) & \forall i=\overline{1,m}\,, \end{cases} \tag{2.8}$$

где $L(x,\lambda)$ — регулярная функция Лагранжа линейной задачи (2.1), имеющая вид

$$L(x,\lambda)=\sum_{j=1}^{n}\sigma_j x_j-\sum_{i=1}^{m}\lambda_i\Big(-\beta_i+\sum_{j=1}^{n}\alpha_{ij}x_j\Big)\,.$$

А, поскольку функция Лагранжа в регулярном случае представима как $L(x,\lambda)=F(x)-\sum\limits_{i=1}^{m}\lambda_i f_i(x)$ (т.е. в формате, не зависящим от того, линейны функции $F(x)$, $f_i(x)$ $\forall i=\overline{1,m}$ или нет), то возможно использовать (2.7) как *определение* вспомогательной функции $U(\tau,x,\lambda)$ и для нелинейной задачи (1.1).

Свойства функций обратных связей в нелинейных задачах

Исследуем условия применимости метода функций обратных связей для решения *нелинейной* задачи вида (1.1) с конечным значением F^* и, быть может, не единственной точкой x^*. Пусть для этой задачи существует регулярная функция Лагранжа $L(x,\lambda) = F(x) - \sum\limits_{i=1}^{m} \lambda_i f_i(x)$, имеющая седловую точку $\{x^*; \lambda^*\}$.

Предположим также, что в рассматриваемой задаче функция Лагранжа регулярна и существуют компактные, с непустой внутренностью множества $\Omega_x \subseteq E_+^n$, $\Omega_\lambda \subseteq E_+^m$ такие, что найдется хотя бы одна пара векторов $x^* \in \Omega_x$ и $\lambda^* \in \Omega_\lambda$, такая, что $L(x^*, \lambda^*) = F^*$. Заметим так же, что функция $L(x,\lambda)$ в сделанных предположениях непрерывна в $\Omega = \Omega_x \times \Omega_\lambda$.

Введем в рассмотрение *функцию обратных связей* $Q(\tau, s)$ определенную $\forall \tau > 0$ и $\forall s \in (0, +\infty)$.

Для упрощения дальнейших рассуждений введем дополнительное ограничение на выбор функции $Q(\tau, s)$, а именно: будем предполагать[2], что функция $\dfrac{\partial P}{\partial s}\big(\tau, s\big)$ имеет вид $\Phi\left(\dfrac{s}{\tau}\right)$, где $\Phi(u)$ является строго монотонной и непрерывно дифференцируемой функцией *одного* аргумента $u = \dfrac{s}{\tau}$. В этом случае $Q(\tau, s) = \tau\Psi(s)$, где $\Psi(s)$ есть функция, обратная функции $\Phi(u)$. Действительно, по определению обратной функции:

$$\frac{\partial P}{\partial s}\big(\tau, Q\big) = s \quad\Longrightarrow\quad \Phi\left(\frac{Q}{\tau}\right) = s \quad\Longrightarrow\quad \frac{Q}{\tau} = \Psi(s) \quad\Longrightarrow\quad Q(\tau, s) = \tau\Psi(s)\,.$$

[2]Выбор вида функции $Q(\tau, s)$, равно как и другие сделанные выше предположения, обусловлены желанием упростить обоснование рассматриваемой схемы. Вопрос о том, насколько это влияет на ее практическую эффективность, в данной статье не рассматривается и требует отдельного исследования.

Пусть функция $\Psi(s)$ обладает следующими свойствами:

3–1°. $\Psi(s)$ строго монотонно возрастает по s
и при этом $\lim\limits_{s\to 0+} \Psi(s) = -\infty$, $\lim\limits_{s\to +\infty} \Psi(s) = +\infty$.

3–2°. В своей области определения функция $\Psi(s)$ непрерывно дифференцируема.

3–3°. $\forall s \in (0, +\infty)$ существует функция $R(\tau, s)$, такая, что

$$R(\tau, s) = \tau \int\limits_{\alpha}^{s} \Psi(u)\, du\,, \tag{3.1}$$

где α есть решение уравнения $\Psi\bigl(\alpha\bigr) = 0$ (в сделанных предположениях такое α существует и единственно) .

Используя формулу (2.7) *как определение*, построим для задачи (1.1) вспомогательную функцию

$$U(\tau, x, \lambda) \;=\; L(x, \lambda) + W(\tau, x, \lambda)\,, \qquad \text{где} \quad W(\tau, x, \lambda) \;=\; -\sum_{j=1}^{n} R(\tau, x_j) + \sum_{i=1}^{m} R(\tau, \lambda_i)\,, \tag{3.2}$$

Свойства функций $R(\tau, s)$, $W(\tau, x, \lambda)$ и $U(\tau, x, \lambda)$ описывают следующие утверждения.

Лемма 3.1. *Функция* $R(\tau, s)$ *при любом фиксированном* $\tau > 0$ *и* $\forall s > 0$:

1) неотрицательна, дважды непрерывно дифференцируема, для нее справедливо предельное равенство: $\lim\limits_{\tau \to 0+} R(\tau, s) = 0 \quad \forall s \in (0, +\infty)$, *причем данный предельный переход монотонный.*

2) кроме того, $\dfrac{\partial^2 R}{\partial s^2} > 0$, *т.е.* $R(\tau, s)$ *строго выпукла по* s*,*

3) имеет при $s = \alpha$ *единственный минимум, в котором* $R\bigl(\tau, \alpha\bigr) = 0$.

Лемма 3.2. *Функция* $W(\tau, x, \lambda)$ *имеет для любого фиксированного* $\tau > 0$:

1) единственный нулевой минимум по $\lambda \in E_+^m$ *в точке с координатами* $\lambda_i = \alpha \ \ \forall i = \overline{1, m}$ *для каждого фиксированного* $x \in \mathrm{int}\Omega_x$.

2) единственный нулевой максимум по $x \in E_+^n$ *в точке с координатами* $x_j = \alpha \ \ \forall j = \overline{1, n}$ *для каждого фиксированного* $\lambda \in \mathrm{int}\Omega_\lambda$.

По построению, множество $\Omega = \Omega_x \times \Omega_\lambda$ есть компакт в $E^n_+ \times E^m_+$. Тогда справедлива

Теорема 3.1. *При фиксированном $\tau > 0$ функция $U(\tau, x, \lambda)$ имеет $\{\,\overline{x}(\tau);\, \overline{\lambda}(\tau)\}$ — единственную седловую точку внутри Ω , где векторы $\overline{x}(\tau)$ и $\overline{\lambda}(\tau)$ — решения системы уравнений*

$$\begin{cases} \operatorname*{grad}_x U(\tau, \overline{x}, \overline{\lambda}) = o, \\ \operatorname*{grad}_\lambda U(\tau, \overline{x}, \overline{\lambda}) = o. \end{cases} \tag{3.3}$$

Заметим, что система (3.3) может быть записана также в виде

$$\begin{cases} \dfrac{\partial L}{\partial x_j} = Q(\tau, x_j) \quad \forall j = \overline{1,n}\,, \\ \dfrac{\partial L}{\partial \lambda_i} = -Q(\tau, \lambda_i) \quad \forall i = \overline{1,m} \end{cases} \quad \text{или} \quad \begin{cases} \dfrac{\partial F}{\partial x_j}(\overline{x}) - \sum\limits_{i=1}^{m} \overline{\lambda}_i \dfrac{\partial f_i}{\partial x_j}(\overline{x}) = Q(\tau, \overline{x}_j) \quad \forall j = \overline{1,n}\,, \\ f_i(\overline{x}) = Q(\tau, \overline{\lambda}_i) \quad \forall i = \overline{1,m}\,, \end{cases} \tag{3.4}$$

оправдывающим использование для функции $Q(\tau, s)$ наименования «*функция обратных связей*».

Иначе говоря, из равенств $\overline{\lambda}_i f_i(\overline{x}) = \overline{\lambda}_i Q(\tau, \overline{\lambda}_i) = \overline{\lambda}_i \tau \Psi(\overline{\lambda}_i) \to 0$ при $\tau \to 0$ следует выполнение условий *дополняющей нежесткости.*

Далее дадим определение: вектор-функцию $\{\overline{x}(\tau); \overline{\lambda}(\tau)\}\ \forall\tau > 0$ будем называть *седловой траекторией* функции $U(\tau, x, \lambda)$. И будем также предполагать, что $\exists\tau_0 > 0$, такое, что $\forall\tau \in (0, \tau_0]$ и

$$\begin{cases} \overline{x}(\tau) \in \Omega_x, \\ \overline{\lambda}(\tau) \in \Omega_\lambda. \end{cases}$$

Теорема 3.2. *На седловой траектории задачи* (1.1)

$$\lim_{\tau\to 0+} U\Big(\tau, \overline{x}(\tau), \overline{\lambda}(\tau)\Big) = F^*, \tag{3.5}$$

а в случае локальной единственности решения задачи (1.1) *справедливы также равенства*

$$\lim_{\tau\to 0+} \overline{x}(\tau) = x^* \qquad \text{\textit{и}} \qquad \lim_{\tau\to 0+} \overline{\lambda}(\tau) = \lambda^* . \tag{3.6}$$

Важное свойство вектор-функций $\{\overline{x}(\tau), \overline{\lambda}(\tau)\}$ описывает

Теорема 3.3. *На седловой траектории вектор-функции* $\{\,\overline{x}(\tau); \overline{\lambda}(\tau)\}$ *непрерывно дифференцируемы* $\forall \tau \in (0, \tau_0]\,.$

Последовательная линейная экстраполяция для функций обратных связей

Непрерывная дифференцируемость вектор-функций $\overline{x}(\tau)$ и $\overline{\lambda}(\tau)$ на седловой траектории для задачи (1.1) позволяет получить оценки x^* и λ^* по схеме, альтернативной непосредственному поиску седловой точки вспомогательной функции (3.1) при малых положительных значениях τ.

Из теоремы 3.3 вытекает, что зависимости $\overline{x}(\tau)$ и $\overline{\lambda}(\tau)$ являются непрерывно дифференцируемыми функциями параметра τ на промежутке $\tau \in (0, \tau_0]$. Тогда на этом промежутке для них существуют тейлоровские аппроксимации

$$\overline{x}(\tau + \Delta) = \overline{x}(\tau) + \frac{d\overline{x}}{d\tau}\Delta + o(\Delta) \qquad \text{и} \qquad \overline{\lambda}(\tau + \Delta) = \overline{\lambda}(\tau) + \frac{d\overline{\lambda}}{d\tau}\Delta + o(\Delta) \ .$$

Если в этих равенствах перейти к пределу при $\Delta \to -\tau + 0$, то получатся следующие оценки для x^* и λ^* :

$$x^+ = \overline{x}(\tau) - \frac{d\overline{x}}{d\tau}\,\tau \qquad \text{и} \qquad \lambda^+ = \overline{\lambda}(\tau) - \frac{d\overline{\lambda}}{d\tau}\,\tau \,. \tag{4.1}$$

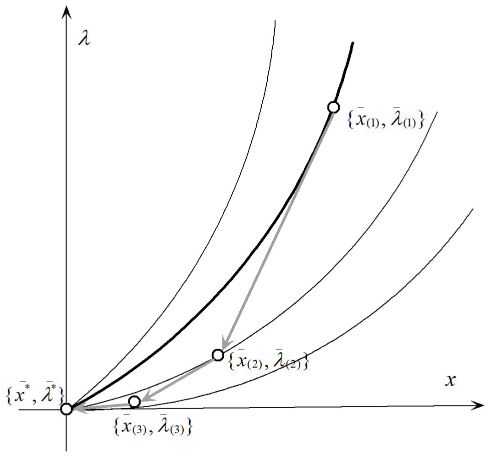

Рис. 4.1. Седловые траектории в последовательной линейной экстраполяции.

При этом значения компонент производных в (4.1) находятся (согласно известной теореме о неявных функциях) из системы линейных уравнений

$$\begin{cases} \sum\limits_{j=1}^{n} \dfrac{\partial^2 U}{\partial x_p \partial x_j} \dfrac{d\overline{x}_j}{d\tau} + \sum\limits_{i=1}^{m} \dfrac{\partial^2 U}{\partial x_p \partial \lambda_i} \dfrac{d\overline{\lambda}_i}{d\tau} = -\dfrac{\partial^2 U}{\partial x_p \partial \tau} \qquad \forall p = \overline{1, n}, \\ \sum\limits_{j=1}^{n} \dfrac{\partial^2 U}{\partial \lambda_q \partial x_j} \dfrac{d\overline{x}_j}{d\tau} + \sum\limits_{i=1}^{m} \dfrac{\partial^2 U}{\partial \lambda_q \partial \lambda_i} \dfrac{d\overline{\lambda}_i}{d\tau} = -\dfrac{\partial^2 U}{\partial \lambda_r \partial \tau} \qquad \forall q = \overline{1, m}. \end{cases} \tag{4.2}$$

которая получается дифференцированием по параметру τ уравнений (3.3) — условий стационарности вспомогательной функции $U(\tau, \overline{x}(\tau), \overline{\lambda}(\tau))$.

Если коэффициент τ достаточно мал по модулю, то в силу теоремы Тейлора в пространстве $E^n \otimes E^m$ норма разности точек $\{x^+, \lambda^+\}$ и $\{x^*, \lambda^*\}$ будет меньше, чем норма разности $\{\overline{x}, \overline{\lambda}\}$ и $\{x^*, \lambda^*\}$. Однако в данной ситуации повторное применение формул (4.1) для повышения точности оказывается неприменимым, поскольку точка $\{x^+, \lambda^+\}$ вообще говоря не будет принадлежать седловой траектории. Преодолеть это затруднение и сделать возможным использование равенств (4.1) для итерационного уточнения решения позволяет следующая схема.

Заменим в функции $W(\tau, x, \lambda)$ скалярный параметр τ на $(n+m)$-компонентный вектор $\vec{\tau}$ с координатным представлением $\| \, \tau_{x1} \, \tau_{x2} \, \ldots \, \tau_{xn} \, \tau_{\lambda 1} \, \tau_{\lambda 2} \, \ldots \, \tau_{\lambda m} \, \|^{\mathrm{T}}$ таким образом, чтобы основные свойства функции $W(\tau, x, \lambda)$, указанные в лемме 3.2, сохранялись.

Теперь путем подбора значений компонент вектора $\vec{\tau}$ обеспечим принадлежность, полученной аппроксимации решения, некоторой новой седловой траектории, сходящейся к $\{x^*, \lambda^*\}$. При такой замене имеем

$$W(\vec{\tau}, x, \lambda) \; = \; -\sum_{j=1}^{n} R(\tau_{xj}, x_j) + \sum_{i=1}^{m} R(\tau_{\lambda i}, \lambda_i) \, ,$$

и возникает новая вспомогательная функция $U\Big(\vec{\tau}, x, \lambda\Big)$, также имеет вид (3.2).

В этом случае оказывается возможным построение процедуры *последовательной линейной экстраполяции* — итеративного улучшения оценок для x^* и λ^* по следующим из (4.1) формулам $\forall t = 0, 1, 2, \ldots$

$$\left\{ \begin{array}{rcl} \overline{x}_{j(t+1)} & = & \overline{x}_{j(t)} - \tau_{xj(t)} \dfrac{d\overline{x}_j}{d\tau_{xj(t)}} \qquad \forall j = \overline{1,n}\,, \\ \\ \overline{\lambda}_{i(t+1)} & = & \overline{\lambda}_{i(t)} - \tau_{\lambda i(t)} \dfrac{d\overline{\lambda}_i}{d\tau_{\lambda i(t)}} \qquad \forall i = \overline{1,m}\,, \end{array} \right. \tag{4.3}$$

где значения производных находятся по формулам (4.2) в точках $\{\, \overline{x}_{(t)}, \overline{\lambda}_{(t)} \}$.

Выбор в формулах (4.3) значений параметров $\{\ \tau_{x1}\, \tau_{x2}\ \ldots\ \tau_{xn}\, \tau_{\lambda 1}\, \tau_{\lambda 2}\ \ldots\ \tau_{\lambda m}\ \}$ делается так, чтобы выполнялись равенства

$$\left\{ \begin{array}{l} \operatorname{grad}_x \widehat{U}\Big(\vec{\tau}_{(t+1)}, \overline{x}_{(t+1)}, \overline{\lambda}_{(t+1)} \Big) = o\,, \\ \operatorname{grad}_\lambda \widehat{U}\Big(\vec{\tau}_{(t+1)}, \overline{x}_{(t+1)}, \overline{\lambda}_{(t+1)} \Big) = o\,. \end{array} \right. \tag{4.4}$$

При этом в силу (3.1) этот пересчет сводится к решению $n+m$ независимых линейных уравнений с одним неизвестным каждое, имеющих вид

$$\begin{cases} \dfrac{\partial L}{\partial x_j}\left(\overline{x}_{(t)}, \overline{\lambda}_{(t)}\right) - \tau_{xj(t+1)}\Psi(\overline{x}_{j(t)}) = 0 \quad \forall j = \overline{1,n}\,, \\ \dfrac{\partial L}{\partial \lambda_i}\left(\overline{x}_{(t)}, \overline{\lambda}_{(t)}\right) - \tau_{\lambda i(t+1)}\Psi(\overline{\lambda}_{i(t)}) = 0 \quad \forall i = \overline{1,m} \end{cases} \quad \text{или} \begin{cases} \dfrac{\partial L}{\partial x_j}\left(\overline{x}_{(t)}, \overline{\lambda}_{(t)}\right) - \tau_{xj(t+1)}\Psi(\overline{x}_{j(t)}) = 0 \quad \forall j = \overline{1,} \\ f_i\left(\overline{x}_{(t)}\right) - \tau_{\lambda i(t+1)}\Psi(\overline{\lambda}_{i(t)}) = 0 \qquad\quad \forall i = \overline{1,} \end{cases}$$

Значения производных в (4.3) для каждого t находятся из (аналогичной (4.2)) системы

$$\begin{cases} \sum\limits_{j=1}^{n} \dfrac{\partial^2 U}{\partial x_p \partial x_j}\dfrac{d\overline{x}_j}{d\tau_{xj(t)}} + \sum\limits_{i=1}^{m} \dfrac{\partial^2 U}{\partial x_p \partial \lambda_i}\dfrac{d\overline{\lambda}_i}{d\tau_{\lambda i(t)}} = -\dfrac{\partial^2 U}{\partial x_p \partial \tau_{xp(t)}} \qquad \forall p = \overline{1,n}, \\ \sum\limits_{j=1}^{n} \dfrac{\partial^2 U}{\partial \lambda_q \partial x_j}\dfrac{d\overline{x}_j}{d\tau_{xj(t)}} + \sum\limits_{i=1}^{m} \dfrac{\partial^2 U}{\partial \lambda_q \partial \lambda_i}\dfrac{d\overline{\lambda}_i}{d\tau_{\lambda i(t)}} = -\dfrac{\partial^2 U}{\partial \lambda_q \partial \tau_{\lambda q(t)}} \qquad \forall q = \overline{1,m}. \end{cases} \tag{4.5}$$

Локальная сходимость описанной процедуры, следует из свойств тейлоровской аппроксимации и очевидной сжимаемости, определяемого формулами (4.3)—(4.4), оператора, при стремящейся к нулю норме вектора $\vec{\tau}$.

Практическое применение процедуры последовательной линейной экстраполяции проиллюстрируем решением следующей задачи.

Задача 1.

Максимизировать $F = -(x_1-1)^2-x_2^2$ при условиях $x_1 \geq 0$, $x_2 \geq 0$ и $x_1+2x_2 \leq 3$, $x_1^2 - x_2 \leq 0$.

Пусть функция $Q(\tau, s) = \dfrac{\tau}{2}\left(s - \dfrac{1}{s}\right)$, тогда соответствующая ей функция $R(\tau, s) = \dfrac{\tau}{2}\left(\dfrac{s^2}{2} - \ln s - \dfrac{1}{2}\right)$.

Введем обозначения

$$\begin{array}{ll} f_1(x) & = -3 + \overline{x}_1 + 2\overline{x}_2 \,, \\ f_2(x) & = \quad x_1^2 - x_2 \,, \\ L(x,\Lambda) & = -(x_1-1)^2 - x_2^2 - \lambda_1(-3 + x_1 + 2x_2) - \lambda_2(x_1^2 - x_2) \,. \end{array}$$

Как было предложено выше, в процедуре последовательной линейной экстраполяции вместо параметра τ используем четырехкомпонентный вектор $\vec{\tau}$ с координатным представлением $\|\ \tau_{x1}\ \tau_{x2}\ \tau_{\lambda1}\ \tau_{\lambda2}\ \|$. Вспомогательная U-функция в рассматриваемой задаче имеет вид

$$U(\tau, x, \Lambda) = -(x_1-1)^2-x_2^2-\lambda_1(-3+x_1+2x_2)-\lambda_2(x_1^2-x_2)-R(\tau_{x1}, x_1)-R(\tau_{x2}, x_2)+R(\tau_{\lambda1}, \lambda_1)+R(\tau_{\lambda2}, \lambda_2)$$

Тогда система уравнений (3.4) – условий стационарности U-функции — для решаемой задачи будет

$$\begin{cases} -3+\overline{x}_1+2\overline{x}_2=\dfrac{\tau}{2}\left(\overline{\lambda}_1-\dfrac{1}{\overline{\lambda}_1}\right), & \overline{x}_1^2-\overline{x}_2=\dfrac{\tau}{2}\left(\overline{\lambda}_2-\dfrac{1}{\overline{\lambda}_2}\right), \\ -2\Big(\overline{x}_1-1\Big)-\overline{\lambda}_1-2\overline{x}_1\cdot\overline{\lambda}_2=\dfrac{\tau}{2}\left(\overline{x}_1-\dfrac{1}{\overline{x}_1}\right), & -2\overline{x}_2-2\overline{\lambda}_1+\overline{\lambda}_2=\dfrac{\tau}{2}\left(\overline{x}_2-\dfrac{1}{\overline{x}_2}\right). \end{cases}$$

За начальное приближение в процедуре последовательной линейной экстраполяции примем решение системы (3.4) с $\tau=0.01$, приведенным в табл. 1.

Решения системы (3.4) при $\tau = 0.01$.

τ	$\overline{x}_1(\tau)$	$\overline{x}_2(\tau)$	$\overline{\lambda}_1(\tau)$	$\overline{\lambda}_2(\tau)$	$F(\overline{x}(\tau))$	$f_1(\overline{x}(\tau))$	$f_2(\overline{x}(\tau))$	$L(\overline{x},\overline{\Lambda})$	$U(\overline{x},\overline{\Lambda})$
10^{-2}	0.59002481	0.35181724	$2.930*10^{-3}$	0.69704208	-0.51620893	-1.70634071	$-3.688*10^{-3}$	-0.50863832	-0.48549662

Таблица 1.

Последовательное уточнение значений прямых и двойственных переменных будем выполнять по формулам (4.3) с использованием решений системы линейных уравнений (4.5), основная матрица которой имеет вид

$$\begin{pmatrix} -\dfrac{\tau}{2}\left(1+\dfrac{1}{x_1^2}\right)-2-2\lambda_2 & 0 & -1 & -2x_1 \\ 0 & -\dfrac{\tau}{2}\left(1+\dfrac{1}{x_2^2}\right)-2 & -2 & 1 \\ -1 & -2 & \dfrac{\tau}{2}\left(1+\dfrac{1}{\lambda_1^2}\right) & 0 \\ -2x_1 & 1 & 0 & \dfrac{\tau}{2}\left(1+\dfrac{1}{\lambda_2^2}\right) \end{pmatrix},$$

в то время как столбец правых частей этой системы будет

$$\begin{pmatrix} -\dfrac{\tau}{2}\left(x_1-\dfrac{1}{x_1}\right) & -\dfrac{\tau}{2}\left(x_2-\dfrac{1}{x_2}\right) & \dfrac{\tau}{2}\left(\lambda_1-\dfrac{1}{\lambda_1}\right) & \dfrac{\tau}{2}\left(\lambda_2-\dfrac{1}{\lambda_2}\right) \end{pmatrix}^{\mathrm{T}}.$$

Пусть в процедуре (4.3) уточненные значения переменных на итерации с номером $t-1 \quad \forall t=1,2,\ldots$ равны соответственно $\overline{x}_{1(t)}, \overline{x}_{2(t)}$ и $\overline{\lambda}_{1(t)}, \overline{\lambda}_{2(t)}$. Тогда значения компонент $\vec{\tau}_{(t)}$, гарантирующих принадлежность точки $\{\overline{x}_{(t)}, \overline{\Lambda}_{(t)}\}$ некоторой другой седловой траектории, в силу (4.4) будут определяться равенствами

$$\left\{\begin{array}{lll} \tau_{x1(t)} = \dfrac{2\Big(-2(\overline{x}_{1(t)}-1)-\overline{\lambda}_{1(t)}-2\overline{x}_{1(t)}\overline{\lambda}_{2(t)}\Big)}{\overline{x}_{1(t)}-\dfrac{1}{\overline{x}_{1(t)}}}, & & \tau_{x2(t)} = \dfrac{2\Big(-2\overline{x}_{2(t)}-2\overline{\lambda}_{1(t)}+\overline{\lambda}_{2(t)}\Big)}{\overline{x}_{2(t)}-\dfrac{1}{\overline{x}_{2(t)}}}, \\ \tau_{\lambda1(t)} = \dfrac{2\Big(-3+\overline{x}_{1(t)}+2\overline{x}_{2(t)}\Big)}{\overline{\lambda}_{1(t)}-\dfrac{1}{\overline{\lambda}_{1(t)}}}, & & \tau_{\lambda2(t)} = \dfrac{2\Big(\overline{x}_{1(t)}^2-\overline{x}_{2(t)}\Big)}{\overline{\lambda}_{2(t)}-\dfrac{1}{\overline{\lambda}_{2(t)}}}. \end{array}\right.$$

Величина погрешности на каждой итерации может быть оценена , во-первых, по значениям двойственных переменных λ_1 и λ_2, точные значения которых в данной задаче находятся аналитически и равны соответственно 0 и λ_2^* – положительному корню уравнения $\lambda_2(\lambda_2 + 1)^2 - 2 = 0$, равному (по формуле Кардано) $\lambda_2^* = \left(\sqrt[3]{28+3\sqrt{87}}-1\right)^2 \Big/ 3\sqrt[3]{28+3\sqrt{87}} \approx 0.69562077$. И, во-вторых, по абсолютной величине разности значений целевой функции задачи F и функции Лагранжа L, (см. табл. 3).

Результаты расчетов выполненных в процессе решения задачи 1 приведены в табл. 2, 3 и 4.

t	$\overline{x}_{1(t)}$	$\overline{x}_{2(t)}$	F	$\overline{\lambda}_{1(t)}$	$\overline{\lambda}_{2(t)}$
0	0.590024813	0.351817238	-0.291855023	0.002930222	0.697042081
1	0.589788382	0.347873253	-0.289289372	$-1.390 \cdot 10^{-5}$	0.695539663
2	0.589754522	0.347810387	-0.289289372	$1.294 \cdot 10^{-9}$	0.695620736
3	0.589754512	0.347810385	-0.289273424	0	0.695620770

Т а б л и ц а 2

t	f_1	f_2	L	$L-F$	U
0	-1.706340711	-0.003687958	-0.284284404	0.007570619	-0.261142713
1	-1.714465112	$-2.291 \cdot 10^{-5}$	-0.289297265	$-7.893 \cdot 10^{-6}$	-0.289593289
2	-1.714624703	$8.824 \cdot 10^{-9}$	-0.289273422	$-3.920 \cdot 10^{-9}$	-0.289273422
3	-1.714624718	$-1.11 \cdot 10^{-15}$	-0.289273424	0	-0.290232652

Т а б л и ц а 3

t	$\tau_{x1(t)}$	$\tau_{x2(t)}$	$\tau_{\lambda 1(t)}$	$\tau_{\lambda 2(t)}$	$\overline{\lambda}_{2(t)} - \lambda_2^*$
0	0.01	0.01	0.01	0.01	0.001421311
1	$9.565 \cdot 10^{-6}$	0.000141717	$-4.767 \cdot 10^{-5}$	$6.176 \cdot 10^{-5}$	$-8.111 \cdot 10^{-5}$
2	$-8.794 \cdot 10^{-9}$	$3.265 \cdot 10^{-8}$	$4.437 \cdot 10^{-9}$	$-2.379 \cdot 10^{-8}$	$-3.373 \cdot 10^{-8}$
3	$-1.41 \cdot 10^{-15}$	0	0	$2.99 \cdot 10^{-15}$	$-4.40 \cdot 10^{-10}$

Т а б л и ц а 4

Замечания о практическом использовании метода функций обратных связей

Оценка эффективности практического применения вычислительных схем подобных методу функций обратных связей представляется предметом самостоятельного исследования. Тем не менее, некоторые соображения можно привести уже здесь.

Во-первых, можно избежать необходимости специального контроля за выполнением условия $s > 0$ при вычислении значений функции обратных связей в пробных точках, и понизить зависимость процесса решения системы (3.4) от выбора начального приближения, если заменить неизвестные в системе (3.4) их абсолютными величинами. Иначе говоря, имеет смысл решать не систему (3.4) в неотрицательном ортанте, а систему

$$\left\{ \begin{array}{lcl} \dfrac{\partial L}{\partial \lambda_i}\Big(|\overline{x}_1|,\, |\overline{x}_2|,\, \ldots |\overline{x}_n|,\, |\overline{\lambda}_1|,\, |\overline{\lambda}_2|,\, \ldots |\overline{\lambda}_m|\Big) & = & -\,Q\big(\,\tau, |\overline{\lambda}_i|\,\big) \qquad \forall i = \overline{1, m}\ , \\ \dfrac{\partial L}{\partial x_j}\Big(|\overline{x}_1|,\, |\overline{x}_2|,\, \ldots |\overline{x}_n|,\, |\overline{\lambda}_1|,\, |\overline{\lambda}_2|,\, \ldots |\overline{\lambda}_m|\Big) & = & Q\big(\,\tau, |\overline{x}_j|\,\big) \qquad \forall j = \overline{1, n}\ , \end{array} \right. \tag{5.1}$$

функции в записи которой являются четными по компонентам векторов $\overline{x}$ и $\overline{\lambda}$. В этом случае процесс решения (5.1) может завершиться в любом ортанте пространства $E^n \otimes E^m$, однако решение системы (3.4) очевидно получится из решения (5.1) при замене компонент $\overline{x}$ и $\overline{\lambda}$ их модулями.

Во-вторых, утверждения §3 в свой совокупности позволяют прийти к заключению, что функция $R(\tau, s)$ при предельном переходе $\tau \to +0$ может рассматриваться, с одной стороны, как барьерная (или квазибарьерная) штрафная функция по ограничению $s < 0$ и, с другой стороны, как регуляризирующее слагаемое для переменной s, препятствующее неограниченному росту значения этой переменной.

Основываясь на этом факте, представляется возможным расширение класса функций обратных связей, путем использования функций $R(\tau, s)$ со штрафующими свойствами внешнего или общего типов, при соблюдении всех прочих, сформулированных выше условий.