SCIREA Journal of Mathematics
http://www.scirea.org/journal/Mathematics
September 12, 2022
Volume 7, Issue 5, October 2022
https://doi.org/10.54647/mathematics11348

Feedback Functions in Problems of Nonlinear Programming

A.E. Umnov, E.A. Umnov.

Russian Federation, 141700 Dolgoprudny, MO, Institutskiy per., 9, Moscow Institute of Physics and Technology(National Research University)

e-mail: mail@umnov.ru

Функции обратной связи в задачах нелинейного программирования

А.Е. Умнов, Е.А. Умнов.

Российская Федерация, 141700 Долгопрудный, МО, Институтский пер., 9,

Московский физико-технический институт (национальный исследовательский университет)

Электронная почта: `mail@umnov.ru`

В статье рассматривается метод решения задачи нелинейного программирования. В этом методе применяются специальные функции обратной связи, описывающие связи между прямыми переменными и множителями Лагранжа. Эти связи аналогичны условиям теоремы Каруша-Куна-Таккера, но не используют неравенств в своей записи. С помощью функций обратной связи строится модифицированная функция Лагранжа, седловые точки которой являются приближенными решениями исходной задачи нелинейного программирования. Дано обоснование этой схемы и продемонстрирован пример ее использования.

Ключевые слова: задача нелинейного программирования. Функции обратной связи. Модифицированная функция Лагранжа. Седловая траектория. Метод последовательной экстраполяции.

1. Введение

В данной статье рассматривается использование функций обратной связи (т.е. функций, реализующих в условиях оптимальности дополнительные соотношения между прямыми переменными и множителями Лагранжа) для решения задач нелинейного программирования вида:

$$\begin{aligned} & F(x) \to \max, \\ & \textit{при условиях}: \quad f_i(x) \le 0 \;\; \forall i = \overline{1,m}, \quad x_j \ge 0 \;\; \forall j = \overline{1,n}, \end{aligned} \tag{1.1}$$

где

- $x = (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in E^n$ $-n$-мерное евклидово пространство,
- функции $F(x)$, $f_i(x) \, \forall i = \overline{1,m}$ определены и дважды непрерывно дифференцируемы в неотрицательном ортанте E^n.

Будем предполагать, что задача (1.1) имеет локальное решение (возможно, не единственное) и обозначать $F(x^*)$ как F^*.

Рассматриваемый метод решения задачи (1.1) основан на тех же идеях, что использовались в [1] для решения задачи линейного программирования. Поэтому в следующем разделе сначала приводится краткое описание этих идей для линейного случая, а затем дается описание схемы их использования при решении нелинейных задач.

2. Метод функций обратной связи для нелинейных задач математического программирования

Пусть функции $F(x), f_i(x) \quad \forall i=\overline{1,m}$ линейны в E_+^n − неотрицательном ортанте E^n. Тогда задача (1.1) имеет вид

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

где $\sigma_j, \beta_i, \alpha_{ij} \ \forall j=\overline{1,n}, \forall i=\overline{1,m}$ − константы.

Двойственная линейная задача в этом случае

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

где $\lambda=(\lambda_1,\lambda_2,\dots,\lambda_m)^{\mathrm{T}}$. Решение (2.2) обозначим как $\lambda^*=(\lambda_1^*,\lambda_2^*,\dots,\lambda_m^*)^{\mathrm{T}}$.

Для решения задачи (2.1) применим специальный вариант метода штрафных функций [2,3,4]. Этот метод заключается в последовательном безусловном поиске экстремумов вспомогательной функции

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

где штрафная функция $P(\tau,s)$ удовлетворяет следующим условиям:

1°. $\tau>0 \quad \lim\limits_{\tau\to+0}P(\tau,s)=\begin{cases}+\infty, & s>0,\\ 0, & s<0\end{cases}$ и этот предельный переход монотонный.

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

3°. $\forall\tau>0 \quad \forall s: \quad \dfrac{\partial P}{\partial s}>0, \quad \dfrac{\partial^2 P}{\partial s^2}>0. \qquad (2.3)$

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

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

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

Пусть имеем: $\tilde{x}(\tau)=\arg\max\limits_{x}A_P(\tau,x)$ и $\hat{\lambda}(\tau)=\arg\min\limits_{\lambda}A_D(\tau,\lambda)$. Предположим далее (теория метода штрафных функций допускает такой вариант), что справедливы соотношения:

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

так же как и

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

Другими словами, для линейных задач (2.1) --- (2.2) компоненты вектора – λ^* это множители Лагранжа задачи (2.1), а компоненты вектора x^* – множители Лагранжа задачи (2.2) [3]. Но для фиксированного $\tau > 0$, вообще говоря,

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

Эти соотношения превращаются в верные равенства только при предельном переходе $\tau \to +0$. Однако формально их можно объединить в систему уравнений, справедливую $\forall\,\tau > 0$. Эта система формулируется в терминах новых неизвестных $\bar{x}(\tau)$ и $\bar{\lambda}(\tau)$ в виде

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

В [1] показано, что (2.4) имеет решения для любых задач (2.1) — (2.2). Для совместных задач (2.1) — (2.2)) справедливо соотношение $F^* = \lim\limits_{\tau\to+0} F\left(\bar{x}(\tau)\right)$. Более того, если x^* и λ^* единственны, то также имеют место равенства:

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

Обе вектор-функции $\bar{x}(\tau)$ и $\bar{\lambda}(\tau)$ могут использоваться как приближенные решения задач (2.1) и (2.2). Заметим, что в системе (2.4) нет явных условий неотрицательности компонент $\bar{x}(\tau)$ и $\bar{\lambda}(\tau)$. Этот факт имеет место, поскольку $\dfrac{\partial P}{\partial s} > 0$ по определению $P(\tau, s)$.

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

Тогда систему (2.4) можно записать в виде

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

Заметим, что вид системы (2.5) оправдывает использование термина *функция обратной связи* для функции $Q(\tau, s)$. Примерами функций обратной связи могут служить:

$$\begin{aligned} &1)\quad Q(\tau, s) = \tau \ln s && \text{для} \quad \frac{\partial P}{\partial s} = e^{\frac{s}{\tau}}, \\ &2)\quad Q(\tau, s) = \frac{\tau}{2}\left(s - \frac{1}{s}\right) && \text{для} \quad \frac{\partial P}{\partial s} = \sqrt{\left(\frac{s}{\tau}\right)^2 + 1} + \frac{s}{\tau}. \end{aligned} \tag{2.6}$$

Для перехода к нелинейному случаю введем новую вспомогательную функцию вида

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

где $L(x,\lambda)=\sum_{j=1}^{n}\sigma_j x_j-\sum_{i=1}^{m}\lambda_i\left(-\beta_i+\sum_{j=1}^{n}\alpha_{ij}x_j\right)$ – регулярная функция Лагранжа задачи (2.1), 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. $\tau>0$. Это верно, поскольку функция $Q(\tau,s)$ неограничена как снизу, так и сверху и $Q(\tau,s)$ строго монотонно возрастает по s при $s\in(0,+\infty)$.

В этом случае решения системы (2.5) (т.е. векторы $\bar{x}(\tau)$ и $\bar{\lambda}(\tau)$) являются стационарными точками функции $U(\tau,x,\lambda)$ по переменным x и λ. Наконец, в терминах L и Q система (2.5) может быть записана как

$$\begin{cases}\dfrac{\partial L}{\partial x_j}= Q\left(\tau,\bar{x}_j(\tau)\right) & \forall j=\overline{1,n},\\ \dfrac{\partial L}{\partial\lambda_i}=-Q\left(\tau,\bar{\lambda}_i(\tau)\right) & \forall i=\overline{1,m}.\end{cases} \tag{2.8}$$

Функция Лагранжа для задачи (2.1) записывается в виде

$$L(x,\lambda)=F(x)-\sum_{i=1}^{m}\lambda_i f_i(x)\,. \tag{2.9}$$

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 (2.7) в качестве определения вспомогательной функции $U(\tau,x,\lambda)$ в нелинейном случае. А также использовать ее стационарные точки в качестве приближенных решений задачи (1.1).

В следующем разделе статьи формулируются и доказываются условия, при которых этот подход оказывается применимым.

3. Свойства функций обратной связи в нелинейных задачах.

Основная задача этого параграфа — выяснить условия применимости системы (2.8) для решения нелинейной задачи (1.1).

Пусть задача (1.1) имеет ограниченые F^* и x^*, где x^* может быть не единственным. Пусть также регулярная функция Лагранжа (2.9) имеет для этой задачи седловую точку $\left\{x^*,\lambda^*\right\}$.

При этом предполагается, что существуют компакты (с непустой внутренностью) $\Theta_x,\ \ \Theta_\lambda$, такие, что имеется хотя бы одна пара векторов $\left\{x^*\in\operatorname{int}\Theta_x,\quad \lambda^*\in\operatorname{int}\Theta_\lambda\right\}$, для которых $L\left(x^*,\ \lambda^*\right)=F^*$.

Мы также будем использовать множества $\Omega_x=\Theta_x\bigcap E_+^n,\quad \Omega_\lambda=\Theta_\lambda\bigcap E_+^m$. Заметим, что при сделанных предположениях функция Лагранжа $L\left(x,\ \lambda\right)$ непрерывна на множестве $\Omega=\Omega_x\times\Omega_\lambda$.

Теперь введем функцию обратной связи $Q(\tau,s)$ для $\tau>0$ и $s\in(0,+\infty)$ в следующем специальном виде: $Q(\tau,s)=\tau\,\Psi(s)$, где $\Psi(s)$ обладает следующими свойствами

3-1. $\Psi(s)$ строго монотонно возрастает по s и

$$\lim_{s\to 0+0}\Psi(s)=-\infty,\qquad \lim_{s\to+\infty}\Psi(s)=+\infty,$$

причем предельные переходы по монотонны.

3-2. $\Psi(s)$ непрерывно дифференцируема.

3-3. Имеется функция $R(\tau,s)$, что

$$R(\tau,s)=\tau\int_{\omega}^{s}\Psi(u)\,du, \tag{3.1}$$

где ω есть решение уравнения $\Psi(\omega)=0$. При сделанных предположениях это решение, очевидно, существует и единственно.

Такой выбор вида функции $Q(\tau,s)$, с одной стороны, значительно упрощает дальнейшие рассуждения, а, с другой стороны, является частным случаем метода построения функции обратной связи, описанного в разделе 2.

Действительно, пусть функция $\frac{\partial P}{\partial s}(\tau,s)$ имеет вид $\Phi\left(\frac{s}{\tau}\right)$, где функция одной переменной $\Phi(u)$ такова, что выполняются все условия (2.3). В этом случае, если $\Psi(u)$ – обратная функция для $\Phi(u)$, то по определению обратной функции имеем

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

Примерами таких функций обратной связи могут слузить функции (2.6).

Используя формулу (2.7) в качестве определения, введем для задачи (1.1) вспомогательную функцию

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

Свойства функций $R(\tau,s)$, $W(\tau,x,\lambda)$ и $U(\tau,x,\lambda)$ описываются следующими утверждениями.

Лемма 3.1. 1) Функция $R(\tau,s)$ неотрицательна, дважды непрерывно дифференцируема при любых фиксированных $\tau>0$ и $\forall s\in(0,+\infty)$.

Для любого фиксированного $s>0$ $\lim_{\tau\to+0}R(\tau,s)=0$. Причем этот переход к пределу является монотонным.

2) Функция $R(\tau,s)$ строго выпукла по s.

3) При $s=\omega$ $R(\tau,s)$ имеет единственный минимум с $R(\tau,\omega)=0$.

Доказательство. 1) В силу (3.1) и свойств интеграла имеем $R(\tau,s)\ge 0$. Функция $R(\tau,s)$ дважды непрерывно дифференцируема, поскольку непрерывно дифференцируема функция $Q(\tau,s)$. Равенство $\lim_{\tau\to+0}R(\tau,s)=0$ и монотонность этого предельного перехода также следуют из (3.1).

2) Функция $Q(\tau,s)$ строго монотонно возрастает по s и непрерывно дифференцируема с $\dfrac{\partial Q}{\partial s}=\dfrac{\partial^2 R}{\partial\, s^2}>0$, следовательно, $R(\tau,s)$ строго выпукла по s.

3) Мы имеем $R(\tau,\omega)=0,\quad \dfrac{\partial R}{\partial s}(\tau,\omega)=0,\quad \dfrac{\partial^2 R}{\partial\, s^2}(\tau,\omega)>0$. Следовательно, $R(\tau,s)$ дважды непрерывно дифференцируемая функция и имеет единственный минимум с нулевым значением.

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

1) строгий нулевой минимум в точке $\lambda\in E_+^m$ с координатами $\lambda_i=\omega\quad \forall i=\overline{1,m}$ для каждого фиксированного $x\in\Omega_x$.

2) строгий нулевой максимум в точке $x\in E_+^n$ с координатами $x_j=\omega\quad \forall j=\overline{1,n}$ для каждого фиксированного $\lambda\in\Omega_\lambda$.

Доказательство. 1) Учтем лемму 3.1 и следующее:

– сепарабельность функции $W(\tau,x,\lambda)$ по компонентам λ,

– равенства $\dfrac{\partial W}{\partial\lambda_i}=\left.\dfrac{\partial R}{\partial\, s}\right|_{s=\omega}=Q(\tau,\omega)=0\quad \forall i=\overline{1,m}$ и

– отношения $\dfrac{\partial^2 W}{\partial\,\lambda_i^2}=\left.\dfrac{\partial^2 R}{\partial\, s^2}\right|_{s=\omega}>0\quad \forall i=\overline{1,m}$.

На основании достаточных условий строгого минимума заключаем, что функция $W(\tau,x,\lambda)$ имеет строгий глобальный минимум в точке $\lambda_i=\omega\quad \forall i=\overline{1,m}$.

2) Доказательство второго утверждения аналогично первому.

В силу сделанных предположений множество $\Omega=\Omega_x\times\Omega_\lambda$ является компактом в $E_+^n\times E_+^m$. Тогда верна следующая теорема.

Теорема 3.1. При фиксированном $\tau>0$ функция $U(\tau,x,\lambda)$ имеет седловую точку $\{\bar{x},\bar{\lambda}\}$ на множестве Ω, где $\bar{x}$ и $\bar{\lambda}$ – решения системы уравнений

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

Доказательство. 1) Сначала покажем, что на множестве Ω функция $U(\tau,x,\lambda)$ имеет максимум по x при каждом фиксированном $\lambda\in\Omega_\lambda$ и минимум по λ при каждом фиксированном $x\in\Omega_x$.

При сделанных предположениях и в силу леммы 3.2 этим свойством будут обладать как функция Лагранжа, так и функция $W(\tau,x,\lambda)$. Согласно (3.2) $U(\tau,x,\lambda)=L(x,\lambda)+W(\tau,x,\lambda)$ на множестве Ω, являющимся компактом. Следовательно, функция $U(\tau,x,\lambda)$ также будет обладать этим свойством.

2) Из теорем о седловой точке [5,6,7,8] следует, что в рассматриваемом случае функция $U(\tau,x,\lambda)$ имеет седловую точку на компакте Ω.

3) Согласно лемме 3.1 функция $R(\tau,s)$ имеет минимум при , $s=\omega$, где $\omega\in(0,+\infty)$. Поэтому седловая точка функции $U(\tau,x,\lambda)$ принадлежит внутренности множества Ω в силу (3.1) и определения множеств Θ_x и Θ_λ , а именно, из строгой выпуклости $R(\tau,s)$ и соотношений $\lim\limits_{s\to 0+0} R(\tau,s)=\lim\limits_{s\to+\infty} R(\tau,s)=+\infty$, верных для любого фиксированного $\tau>0$.

4) Наконец, из того, что функция $U(\tau,x,\lambda)$ дважды непрерывно дифференцируема, следует, что эта седловая точка является стационарной и, следовательно, определяется системой уравнений (3.3). Таким образом, мы приходим к выводу, что утверждение теоремы верно.

Отметим, что систему (3.3) также можно записать в виде

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

Здесь можно также отметить следующее:

1. Система (3.4) оправдывает для $Q(\tau,s)$ название *функция обратной связи* в нелинейном случае.
2. Система (3.4) подобна условиям теоремы Каруша-Куна-Таккера, но не содержит явно условий неотрицательности множителей Лагранжа. Эти условия обеспечиваются свойствами функции $Q\left(\tau,\bar{\lambda}_i\right)$.
3. Функцию $U(\tau,x,\lambda)$ также можно рассматривать как результат специальной модификации функции Лагранжа. Различные виды таких модификаций описаны в большом количестве монографий и статей, например, в [9].
4. С помощью (3.4) легко проверить выполнение условий дополняющей нежесткости. Действительно,

$$\bar{\lambda}_i f_i\left(\bar{x}\right)=\bar{\lambda}_i Q\left(\tau,\bar{\lambda}_i\right)=\tau\bar{\lambda}_i\Psi\left(\bar{\lambda}_i\right)\to 0 \quad \forall i=\overline{1,m}, \qquad \text{при} \quad \tau\to+0 .$$

Теорема 3.2. 1) Для вспомогательной функции $U\left(\tau,x,\lambda\right)$ имеет место равенство

$$\lim_{\tau\to+0} U\left(\tau,x,\lambda\right)=L\left(x,\lambda\right) \quad \forall\{x,\lambda\}\in\Omega$$

2) При сделанных выше предположениях выполнено равенство

$$\lim_{\tau\to+0} U\left(\tau,\bar{x}(\tau),\bar{\lambda}(\tau)\right)=F^* . \tag{3.5}$$

3) В случае локальной единственности решения задачи (1.1) выполняется равенство

$$\lim_{\tau\to+0}\bar{x}(\tau)=x^* . \tag{3.6}$$

Доказательство.

1) Оценим разницу между значениями вспомогательной функции U и функции Лагранжа L. Для точки $\{x,\lambda\}$ в силу леммы 3.1 и (3.2) имеем

$$U(\tau,x,\lambda)=L(x,\lambda)-\sum_{j=1}^{n}R\big(\tau,x_j\big)+\sum_{i=1}^{m}R\big(\tau,\lambda_i\big)\geq L(x,\lambda)-\sum_{j=1}^{n}R\big(\tau,x_j\big)\geq L(x,\lambda)-n\max_{\substack{x\in\Omega_x\\ j=\overline{1,n}}}R\big(\tau,x_j\big)=$$

$$=L(x,\lambda)+n\min_{\substack{x\in\Omega_x\\ j=\overline{1,n}}}R\big(\tau,x_j\big).$$

Аналогично

$$U(\tau,x,\lambda)=L(x,\lambda)-\sum_{j=1}^{n}R\big(\tau,x_j\big)+\sum_{i=1}^{m}R\big(\tau,\lambda_i\big)\leq L(x,\lambda)+\sum_{i=1}^{m}R\big(\tau,\lambda_i\big)\leq L(x,\lambda)+m\max_{\substack{\lambda\in\Omega_\lambda\\ i=\overline{1,m}}}R\big(\tau,\lambda_i\big)=$$

$$=L(x,\lambda)-m\min_{\substack{\lambda\in\Omega_\lambda\\ i=\overline{1,m}}}R\big(\tau,\lambda_i\big).$$

Как было показано ранее, для $\tau>0$ и $s\in(0,+\infty)$ имеем

$$R(\tau,s)\geq 0,\quad \lim_{\tau\to+0}R(\tau,s)=0,\quad \min_{s\in(0,+\infty)}R(\tau,s)=0\ .$$

Тогда, например, согласно лемме 5.1 [6], верны соотношения

$$\lim_{\tau\to+0}\min_{s\in(0,+\infty)}R(\tau,s)=\min_{s\in(0,+\infty)}\lim_{\tau\to+0}R(\tau,s)=0$$

Это дает

$$\begin{cases}\lim\limits_{\tau\to+0}\min\limits_{\substack{x\in\Omega_x\\ j=\overline{1,n}}}R\big(\tau,x_j\big)=\min\limits_{\substack{x\in\Omega_x\\ j=\overline{1,n}}}\lim\limits_{\tau\to+0}R\big(\tau,x_j\big)=0,\\ \lim\limits_{\tau\to+0}\min\limits_{\substack{\lambda\in\Omega_{\lambda x}\\ i=\overline{1,m}}}R\big(\tau,\lambda_i\big)=\min\limits_{\substack{\lambda\in\Omega_{\lambda x}\\ i=\overline{1,m}}}\lim\limits_{\tau\to+0}R\big(\tau,\lambda_i\big)=0.\end{cases}\tag{3.7}$$

Таким образом, получаем оценки

$$L(x,\lambda)+n\min_{\substack{x\in\Omega_x\\ j=\overline{1,n}}}R\big(\tau,x_j\big)\leq U(\tau,x,\lambda)\leq L(x,\lambda)-m\min_{\substack{\lambda\in\Omega_\lambda\\ i=\overline{1,m}}}R\big(\tau,\lambda_i\big).$$

Наконец, используя (3.7), приходим к выводу, что

$$\lim_{\tau\to+0}U\big(\tau,x,\lambda\big)=L\big(x,\lambda\big)\quad \forall\{x,\lambda\}\in\Omega\,.$$

2) Пусть $\tau_0>0$ таково, что $\forall\,\tau\in(0,\tau_0)$ $\begin{cases}\bar{x}(\tau)\in\Omega_x,\\ \bar{\lambda}(\tau)\in\Omega_\lambda.\end{cases}$. Тогда в силу сделанных предположений имеем $\max\limits_{x\in\Omega_x}\min\limits_{\lambda\in\Omega_\lambda}L(x,\lambda)=F^*$. С другой стороны, используя теорему 3.1 и рассуждая как в 1), получаем

$$\max_{x\in\Omega_x}\min_{\lambda\in\Omega_\lambda}L(x,\lambda)=\max_{x\in\Omega_x}\min_{\lambda\in\Omega_\lambda}\lim_{\tau\to+0}U\big(\tau,x,\lambda\big)=\lim_{\tau\to+0}\max_{x\in\Omega_x}\min_{\lambda\in\Omega_\lambda}U\big(\tau,x,\lambda\big)=\lim_{\tau\to+0}U\big(\tau,\bar{x}(\tau),\bar{\lambda}(\tau)\big).$$

Это дает $\lim\limits_{\tau\to+0}U\big(\tau,\bar{x}(\tau),\bar{\lambda}(\tau)\big)=F^*$.

3) $\{\bar{x}(\tau),\bar{\lambda}(\tau)\}$ принадлежит Ω, ограниченному как компакт. Тогда по теореме Больцано-Вейерштрасса $\{\bar{x}(\tau),\bar{\lambda}(\tau)\}$ имеет предельную точку $\{\bar{\bar{x}},\bar{\bar{\lambda}}\}$ при $\tau\to+0$ для которой $L(\bar{\bar{x}},\bar{\bar{\lambda}})=F^*=L(x^*,\lambda^*)$. В силу непрерывности функции Лагранжа $L(x,\lambda)$, если решение задачи (1.1) единственно, то $x^*=\bar{\bar{x}}=\lim\limits_{\tau\to+0}\bar{x}(\tau)$.

Заметим также, что из теоремы 3.2 вытекает следующее.

Пусть некоторое необходимое (и/или достаточное) условие локальной оптимальности для задачи (1.1) сформулировано в терминах функции Лагранжа. Тогда это условие можно также сформулировать в терминах вспомогательной функции $U(\tau, x, \lambda)$.

Назовем вектор-функцию $\{\bar{x}(\tau), \bar{\lambda}(\tau)\}$ $\forall \tau \in (0, \tau_0)$ *седловой траекторией* задачи (1.1). Важное свойство вектор-функции $\{\bar{x}(\tau), \bar{\lambda}(\tau)\}$ описывается следующей теоремой.

Теорема 3.3. На седловой траектории вектор-функция $\{\bar{x}(\tau), \bar{\lambda}(\tau)\}$ непрерывно дифференцируема $\forall \tau \in (0, \tau_0)$.

Доказательство.

1) Покажем сначала, что определитель матрицы Якоби для системы (3.3) не равен нулю. Эта матрица имеет вид

$$H = \left(\begin{array}{c|c} \dfrac{\partial^2 U}{\partial x_j \partial x_k} \ \forall j,k = \overline{1,n} & \dfrac{\partial^2 U}{\partial x_j \partial \lambda_q} \quad \forall j = \overline{1,n}\,, \forall q = \overline{1,m} \\ \hline \dfrac{\partial^2 U}{\partial \lambda_i \partial x_k} \quad \forall i = \overline{1,m}\,, \forall k = \overline{1,n} & \dfrac{\partial^2 U}{\partial \lambda_i \partial \lambda_q} \ \forall i,q = \overline{1,m} \end{array} \right).$$

В силу равенств (2.9), (3.1) и (3.2) имеем

$$\frac{\partial^2 U}{\partial x_j \partial x_k} = A_{jk} = \frac{\partial^2 L}{\partial x_j \partial x_k} - \delta_{jk} \frac{\partial Q}{\partial s}(x_k),$$
$$\frac{\partial^2 U}{\partial x_j \partial \lambda_q} = B_{jq} = -\frac{\partial f_q}{\partial x_j},$$
$$\frac{\partial^2 U}{\partial \lambda_i \partial \lambda_q} = C_{iq} = \delta_{iq} \frac{\partial Q}{\partial s}(\lambda_q),$$

где δ_{ij} − символ Кронекера, т.е. $\delta_{ij} = \begin{cases} 1, \text{if } i = j, \\ 0, \text{if } i \neq j. \end{cases}$ Или, в блочной форме, $H = \left(\begin{array}{c|c} A & B \\ \hline B^{\mathrm{T}} & C \end{array} \right)$.

2) Если $\tau \in (0, \tau_0)$, то определитель подматрицы A ненулевой. Действительно, квадратная матрица A есть сумма матрицы $\nabla^2 L$ и диагональной матрицы с элементами $-\delta_{jk} \dfrac{\partial Q}{\partial s}(x_k)$.

Первое слагаемое является отрицательно полуопределенным в Ω в силу предположения о свойствах функции Лагранжа.

Второе слагаемое строго отрицательно определено, так как $\dfrac{\partial Q}{\partial s} = \dfrac{\partial^2 R}{\partial s^2} > 0$ в силу леммы 3.1. Тогда полная матрица A будет строго отрицательно определена и по критерию Сильвестра $\det A \neq 0$. В этом случае также существует A^{-1}.

Аналогичные рассуждения показывают, что матрица C положительно определена и $\det C > 0$.

3) Используем формулу Шура $\det H = \det A \cdot \det\left(C - B^{\mathrm{T}} A^{-1} B\right)$. Из теоремы Бине-Коши [10] следует, что при любом ранге матрицы B матрица $B^{\mathrm{T}} A^{-1} B$ либо

отрицательно определена, либо отрицательно полуопределена. Тогда матрица $C - B^{\mathrm{T}} A^{-1} B$ положительно определена и $\det\left(C - B^{\mathrm{T}} A^{-1} B\right) \neq 0$.

Наконец, по формуле Шура $\det H = \det A \cdot \det\left(C - B^{\mathrm{T}} A^{-1} B\right) \neq 0$.

4) Пусть $\tau \in (0, \tau_0)$. Рассмотрим систему уравнений (3.3) как неявно задающую вектор-функцию $\left\{\bar{x}(\tau), \bar{\lambda}(\tau)\right\}$. Для фиксированного $\tau > 0$ имеем:

   1. Левые части (3.3) — непрерывно дифференцируемые функции.
   2. Определитель матрицы Якоби для системы (3.3) (или, что то же самое, гессиан вспомогательной функции U) не равен нулю.

   Применим к системе (3.3) теорему о неявной функции [11, 12]. Из этой теоремы следует существование непрерывно дифференцируемой вектор-функции $\left\{\bar{x}(\tau), \bar{\lambda}(\tau)\right\}$ при фиксированном $\tau \in (0, \tau_0)$.

4. Последовательная линейная экстраполяция для функций обратной связи.

Согласно теореме 3.3 вектор-функции $\left\{\bar{x}(\tau), \bar{\lambda}(\tau)\right\}$ непрерывны и дифференцируемы по τ на седловой траектории для задачи (1.1). Это позволяет оценить x^* и λ^* без прямого решения (3.4) при малых $\tau > 0$. Предлагаемый метод основан на тейлоровских аппроксимациях $\left\{\bar{x}(\tau), \bar{\lambda}(\tau)\right\}$.

Пусть $\tau \in (0, \tau_0)$. Для достаточно малых $\Delta > 0$ эти приближения равны

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

Переходя к пределу при $\Delta \to -\tau + 0$, получаем следующие оценки для x^* и λ^*:

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

Значения компонент производных в (4.1) находятся (согласно теореме о неявных функциях) из системы линейных уравнений

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

Уравнения (4.2) получаются дифференцированием уравнений системы (3.3) по переменной τ.

Пусть коэффициент $\tau > 0$ достаточно мал по модулю. По свойствам тейлоровской аппроксимации норма разности точек $\left\{x^+, \lambda^+\right\}$ и $\left\{x^*, \lambda^*\right\}$ будет меньше нормы разности $\left\{\bar{x}(\tau), \bar{\lambda}(\tau)\right\}$ и $\left\{x^*, \lambda^*\right\}$. Если достигнутой точности недостаточно, можно повторить расчеты (4.1)-(4.2) в точке $\left\{x^+, \lambda^+\right\}$ при условии следующей модификации.

Непосредственное повторное использование формул (4.1)-(4.2) невозможно. Действительно, точка $\left\{x^{+},\lambda^{+}\right\}$, вообще говоря, не принадлежит седловой траектории.

Чтобы преодолеть это затруднение, изменим функцию $W(\tau,x,\lambda)$ так, чтобы точка $\left\{x^{+},\lambda^{+}\right\}$ принадлежала какой-то новой седловой траектории. Эта новая траектория должна иметь те же свойства, что и исходная. Рис. 1 иллюстрирует эту идею.

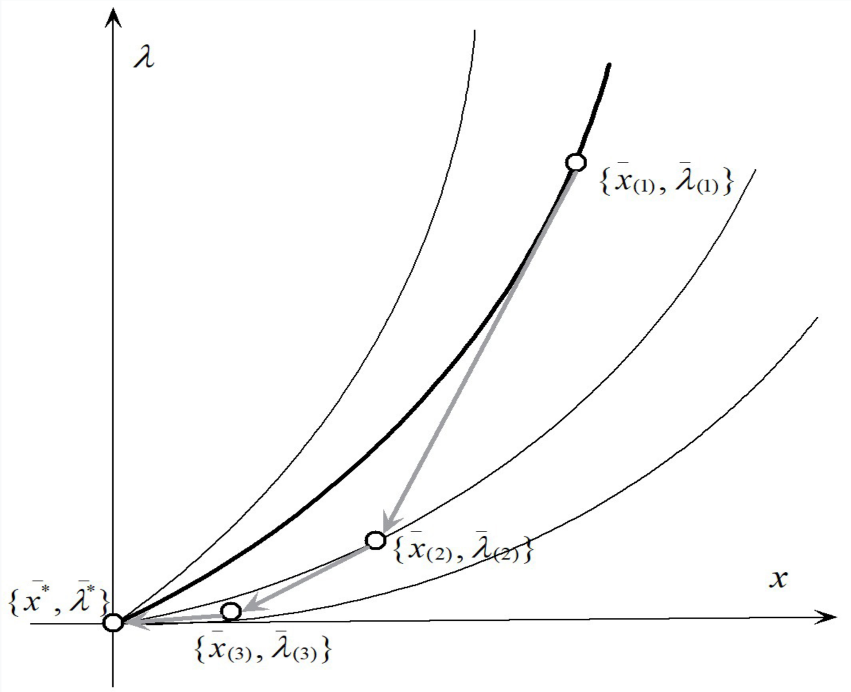

Рис. 1. Седловые траектории при последовательной линейной экстраполяции.

Предлагаемая модификация состоит в следующем.

Заменим скалярный параметр τ в функции $W(\tau,x,\lambda)$ на $(n+m)$-компонентный вектор $\vec{\tau}$. Этот вектор имеет компоненты $\tau_{(t)xj}$ $\forall j=\overline{1,n}$ и $\tau_{(t)\lambda i}$ $\forall i=\overline{1,m}$, где $t=1,2,3,\ldots$ – номер шага экстраполяции.

Все $\tau_{(t)xj}$ $\forall j=\overline{1,n}$ и $\tau_{(t)\lambda i}$ $\forall i=\overline{1,m}$ вводятся аналогично τ в § 3 . Тогда основные свойства

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

такие же, как указано в лемме 3. Вспомогательная функция $U\left(\vec{\tau},x,\lambda\right)$ определяется, как и выше, формулой (3.2).

Пусть $\left\{\bar{x}_{(t)},\bar{\lambda}_{(t)}\right\}$ – текущее приближение на шаге $t=1,2,3,\ldots$ с вектором $\vec{\tau}_{(t)}$. Приближение на следующем шаге будет определяться

$$\begin{cases} \bar{x}_{(t+1)j} = \bar{x}_{(t)j} - \dfrac{d\bar{x}_j}{d\tau_{(t)xj}}\tau_{(t)xj} & j = \overline{1,n}, \\ \bar{\lambda}_{(t+1)i} = \bar{\lambda}_{(t)i} - \dfrac{d\bar{\lambda}_i}{d\tau_{(t)xi}}\tau_{(t)xi} & i = \overline{1,m}. \end{cases} \tag{4.3}$$

Значения производных в (4.3) можно найти с помощью следующей системы линейных уравнений.

$$\begin{cases} \sum\limits_{j=1}^{n}\dfrac{\partial^2 U}{\partial x_p \partial x_j}\cdot\dfrac{d\bar{x}_j}{d\tau_{(t)xj}} + \sum\limits_{i=1}^{m}\dfrac{\partial^2 U}{\partial x_p \partial \lambda_i}\cdot\dfrac{d\bar{\lambda}_i}{d\tau_{(t)\lambda i}} = -\dfrac{\partial^2 U}{\partial x_p \partial \tau_{(t)xp}} & \forall p = \overline{1,n}\ , \\ \sum\limits_{j=1}^{n}\dfrac{\partial^2 U}{\partial \lambda_q \partial x_j}\cdot\dfrac{d\bar{x}_j}{d\tau_{(t)xj}} + \sum\limits_{i=1}^{m}\dfrac{\partial^2 U}{\partial \lambda_q \partial \lambda_i}\cdot\dfrac{d\bar{\lambda}_i}{d\tau_{(t)\lambda i}} = -\dfrac{\partial^2 U}{\partial \lambda_q \partial \tau_{(t)\lambda q}} & \forall q = \overline{1,m}\ . \end{cases} \tag{4.4}$$

Все частные производные рассчитываются при $\{\bar{x}_{(t)}, \bar{\lambda}_{(t)}\}$ и $\vec{\tau}_{(t)}$.

Наконец, компоненты $\vec{\tau}_{(t+1)}$ могут быть найдены из

$$\begin{cases} \underset{x}{\operatorname{grad}} U\left(\vec{\tau}_{(t+1)}, \bar{x}_{(t+1)}, \bar{x}_{(t+1)}\right) = o, \\ \underset{x}{\operatorname{grad}} U\left(\vec{\tau}_{(t+1)}, \bar{x}_{(t+1)}, \bar{x}_{(t+1)}\right) = o. \end{cases} \tag{4.5}$$

При использовании (3.1) и (3.2) система (4.5) принимает вид

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

Обратите внимание, что последняя система представляет собой систему линейных уравнений, каждое из которых имеет только одно неизвестное. *Локальная* сходимость описанной процедуры следует из сжимаемости оператора, определяемого формулами (4.3)–(4.4)–(4.5). Действительно, в силу свойств тейлоровского приближения $\lim\limits_{t\to\infty}\left\|\vec{\tau}_{(t)}\right\| = 0$.

Теперь проиллюстрируем практическое использование метода функций обратной связи совместно с процедурой линейной последовательной экстраполяции для решения следующей задачи нелинейного программирования.

Задача 1. *Максимизировать в* E^2 $\quad F(x) = -(x_1 - 1)^2 - x_2^2$

при условиях $x_1 \ge 0, x_2 \ge 0$ *и* $x_1 + 2x_2 \le 3, \quad x_1^2 - x_2 \le 0$..

Пусть функция обратной связи $Q(\tau, s) = \dfrac{\tau}{2}\left(s - \dfrac{1}{s}\right)$. Заметим, что в этом случае

$$R(\tau, s) = \frac{\tau}{2}\left(\frac{s^2}{2} - \ln s - \frac{1}{2}\right) \quad \text{и} \quad \frac{\partial P}{\partial s} = \sqrt{\left(\frac{s}{\tau}\right)^2 + 1} + \frac{s}{\tau}\ .$$

Введем обозначения

$$f_1(x) = -3 + x_1 + 2x_2,$$
$$f_2(x) = \ \ x_1^2 - x_2,$$
$$L(x,\lambda) = -(x_1-1)^2 - x_2^2 - \lambda_1\left(-3 + x_1 + 2x_2\right) - \lambda_2\left(x_1^2 - x_2\right).$$

В процедуре последовательной линейной экстраполяции вместо τ используем четырехкомпонентный вектор $\vec{\tau}$ с координатным столбцом $\left(\tau_{x1}, \tau_{x2}, \tau_{\lambda 1}, \tau_{\lambda 2}\right)^{\mathrm{T}}$.

Вспомогательная U-функция в задаче 1 имеет вид

$$U(\tau,x,\lambda) =$$
$$= -(x_1-1)^2 - x_2^2 - \lambda_1\left(-3 + x_1 + 2x_2\right) - \lambda_2\left(x_1^2 - x_2\right) - R\left(\tau_{x1}, x_1\right) - R\left(\tau_{x2}, x_2\right) + R\left(\tau_{\lambda 1}, \lambda_1\right) + R\left(\tau_{\lambda 2}, \lambda_2\right).$$

Тогда система уравнений (3.4) (условия стационарности U-функции) для будет

$$\begin{cases} -3 + \bar{x}_1 + 2\bar{x}_2 = \dfrac{\tau_{\lambda 1}}{2}\left(\bar{\lambda}_1 - \dfrac{1}{\bar{\lambda}_1}\right), \\ \bar{x}_1^2 - \bar{x}_2 = \dfrac{\tau_{\lambda 2}}{2}\left(\bar{\lambda}_2 - \dfrac{1}{\bar{\lambda}_2}\right), \\ -2(\bar{x}_1 - 1) - \bar{\lambda}_1 - 2\bar{x}_1\bar{\lambda}_2 = \dfrac{\tau_{x1}}{2}\left(\bar{x}_1 - \dfrac{1}{\bar{x}_1}\right), \\ -2\bar{x}_2 - \overline{2\lambda}_1 + \bar{\lambda}_2 = \dfrac{\tau_{x2}}{2}\left(\bar{x}_2 - \dfrac{1}{\bar{x}_2}\right). \end{cases} \tag{4.6}$$

В качестве начального приближения в процедуре последовальной линейной экстраполяции возьмем решение системы (4.6) с $\tau = 0.01$. Это приближение приведено в табл.1А и табл. 1Б.

Таблица 1А. Решение системы (4.6) при $\tau = 0.01$.

$\bar{x}_1(\tau)$	$\bar{x}_2(\tau)$	$\bar{\lambda}_1(\tau)$	$\bar{\lambda}_2(\tau)$	$F\left(\bar{x}(\tau)\right)$
0.59002481	0.35181724	$2.930\cdot 10^{-3}$	0.69704208	-0.51620893

Таблица 1Б. Решение системы (4.6) при $\tau = 0.01$.

$f_1\left(\bar{x}(\tau)\right)$	$f_2\left(\bar{x}(\tau)\right)$	$L\left(\bar{x}(\tau), \bar{\lambda}(\tau)\right)$	$U\left(\bar{x}(\tau), \bar{\lambda}(\tau)\right)$
1.70634071	$-3.688\cdot 10^{3}$	0.50863832	0.48549662

Матрица системы (4.5) есть

$$\begin{pmatrix} -\dfrac{\tau}{2}\left(1+\dfrac{1}{\bar{x}_1^2}\right)-2-2\bar{\lambda}_2 & 0 & -1 & -2\bar{x}_1 \\ 0 & -\dfrac{\tau}{2}\left(1+\dfrac{1}{\bar{x}_1^2}\right)-2 & -2 & 1 \\ -1 & -2 & \dfrac{\tau}{2}\left(1+\dfrac{1}{\bar{\lambda}_1^2}\right) & 0 \\ -2\bar{x}_1 & 1 & 0 & \dfrac{\tau}{2}\left(1+\dfrac{1}{\bar{\lambda}_2^2}\right) \end{pmatrix},$$

в то время как *RHS*-столбец этой системы

$$\begin{pmatrix} -\dfrac{\tau}{2}\left(\bar{x}_1-\dfrac{1}{\bar{x}_1}\right) & -\dfrac{\tau}{2}\left(\bar{x}_2-\dfrac{1}{\bar{x}_2}\right) & \dfrac{\tau}{2}\left(\bar{\lambda}_1-\dfrac{1}{\bar{\lambda}_1}\right) & \dfrac{\tau}{2}\left(\bar{\lambda}_2-\dfrac{1}{\bar{\lambda}_2}\right) \end{pmatrix}^{\mathrm{T}}.$$

Пусть компоненты $\bar{x}$ и $\bar{\lambda}$ и в итерации с номером t $\forall t=1,2,3,\ldots$ равны $\bar{x}_{1(t)},\bar{x}_{2(t)}$ и $\bar{\lambda}_{1(t)},\bar{\lambda}_{2(t)}$. Тогда значения компонент $\vec{\tau}_{(t+1)}$, обеспечивающие принадлежность точки $\left\{\bar{x}_{(t+1)},\bar{\lambda}_{(t+1)}\right\}$ какой-либо другой седловой траектории, следуют из (4.4)

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

Величина погрешности может быть оценена на каждой итерации двумя различными способами.

Во-первых, с помощью $\lambda_1^*=0$ и того, что λ_2^* является единственным положительным корнем уравнения $\lambda(\lambda+1)^2-2=0$. По формуле Кардано имеем

$$\lambda_2^*=\frac{\left(\sqrt[3]{28+3\sqrt{87}}-1\right)^2}{3\sqrt[3]{28+3\sqrt{87}}}\approx 0.69562077.$$

Во-вторых, мы можем получить эту оценку, сравнив $\left|L-F\right|$ с нулем. Результаты выполненных расчетов приведены в табл.~2, 3 и 4.

Таблица 2. Задача 1 (линейная экстраполяция).

t	$\bar{x}_{1(t)}(\tau)$	$\bar{x}_{2(t)}(\tau)$	$F\left(\bar{x}_{(t)}(\tau)\right)$	$\bar{\lambda}_{1(t)}(\tau)$	$\bar{\lambda}_{2(t)}(\tau)$
1	0.590024813	0.351817238	-0.291855023	0.002930222	0.697042081
2	0.589788382	0.347873253	-0.289289372	$-1.390\cdot10^{-5}$	0.695539663
3	0.589754522	0.347810387	-0.289289372	$1.294\cdot10^{-8}$	0.695620736
4	0.589754512	0.347810385	-0.289273424	0	0.695620770

Таблица 3. Задача 1 (линейная экстраполяция).

t	$f_1(\bar{x}_{(t)}(\tau))$	$f_2(\bar{x}_{(t)}(\tau))$	L	L-F	U
1	-1.706340711	-0.003687958	-0.284284404	0.007570619	-0.261142713
2	-1.714465112	$-2.291\cdot10^{-5}$	-0.289297265	$-7.893\cdot10^{-6}$	-0.289593289
3	-1.714624703	$8.824\cdot10^{-9}$	-0.289273422	$-3.920\cdot10^{-9}$	-0.289273422
4	-1.714624718	$-1.11\cdot10^{-15}$	-0.289273424	0	-0.290232652

Таблица 4. Задача 1 (линейная экстраполяция).

t	$\tau_{x1(t)}$	$\tau_{x2(t)}$	$\tau_{\lambda1(t)}$	$\tau_{\lambda2(t)}$	$\bar{\lambda}_{2(t)}-\lambda_2^*$
1	0.01	0.01	0.01	0.01	0.001421311
2	$9.565\cdot10^{-6}$	0.000141717	$-4.767\cdot10^{-5}$	$6.176\cdot10^{-5}$	$-8.111\cdot10^{-5}$
3	$-8.794\cdot10^{-9}$	$3.265\cdot10^{-8}$	$4.437\cdot10^{-9}$	$-2.379\cdot10^{-8}$	$-3.373\cdot10^{-8}$
4	$-1.41\cdot10^{-15}$	0	0	$2.99\cdot10^{-15}$	$-4.40\cdot10^{-10}$

5. Замечание о практическом использовании метода функций обратной связи.

По мнению авторов, эффективность описанного метода является предметом отдельного исследования. Однако здесь уже можно сделать некоторые выводы.

На практике часто бывает так, что некоторые пробные точки численной процедуры решения системы (3.4) оказываются вне положительного ортанта. Избежать необходимости специального контроля за выполнением условия $s>0$ можно следующим образом.

Заменим неизвестные $\bar{x}_j\ \forall j=\overline{1,n}$ и $\bar{\lambda}_i\ \forall i=\overline{1,m}$ в системе (3.4) их модулями $|\bar{x}_j|\ \forall j=\overline{1,n}$ и $|\bar{\lambda}_i|\ \forall i=\overline{1,m}$. Другими словами, вместо (3.4) будем решать систему

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

без каких-либо дополнительных ограничений.

Решения системы (5.1) могут не принадлежать положительному ортанту пространства $E^n \times E^m$. Однако очевидно, что их абсолютные значения являются положительными решениями системы (3.4).

Кроме того, такая замена снижает зависимость вычислительной процедуры от выбора начальной точки.

Литература

[1] *А.Е. Умнов, Е.А. Умнов*. Использование функций обратной связи в задачах линейного программирования // Вычислительная математика и математическая физика. 59, № 10, стр. 1626–1638. 2019.

[2] *Dimitri P. Bertsekas*. Nonlinear Programming. Athena Scientific, Belmont MA, 2016.

[3] *Fiacco~A.V., McCormic~G.P.* Nonlinear Programming. Sequential Unconstrained Minimization Techniques. N.Y.~: John Wiley and Sons, 1968.

[4] *Умнов А.Е.* Многошаговая линейная экстраполяция в методе штрафных функций // Вычислительная математика и математическая физика. 1974. Т.~14, №~6. - С.~1451-1463.

[5] *Bazaraa M.S., Sherali H.D., Shetty C.M.* Nonlinear programming: Theory and Algorithms. New Jersey, John Wiley \& Sons, Inc., 2006.

[6] *Сухарев А.Г., Тимохов А.В., Федоров В.В.* Курс методов оптимизации. М.~: Физматлит, 2011.

[7] *Jorge Nocedal, Stephen J. Wright.* Numerical Optimization. Springer-Verlag, Berlin, New-York, 2006.

[8] *Гермейер Ю.Б.* Введение в теорию исследования операций. Москва.~:Наука, 1971.

[9] *Скарин В.Д.* Аппроксимация и регуляризующие свойства штрафных функций и функций Лагранжа в математическом программировании. Автореферат диссертации на соискание ученой степени кандидата технических наук, ИММ УрО РАН. Екатеринбург, 2010.

[10] *Фаддеев Д.К.* Лекции по алгебре. Москва.~:Наука, 1984.

[11] *Steven G. Krantz, Harold R. Parks.* The Implicit Function Theorem: History, Theory and Application. Birkhauser, Boston, Basel, Berlin, 2013.

[12] *Кудрявцев Л.Д.* Курс математического анализа. Том.~2. Москва.~: Высшая школа, 1981.