Планы ответов на вопросы экзаменационных билетов госэкзамена
по курсу **ОПТИМИЗАЦИЯ И ЧИСЛЕННЫЕ МЕТОДЫ**,
лектор проф. М. М. Потапов

Вопрос:

4. Симплекс-метод для канонической задачи линейного программирования: идея метода и ее реализация, выбор стартовой угловой точки.

План ответа.

Общая задача линейного программирования:

$$J(u) = \langle c, u \rangle \to \inf, \quad u \in U = \{u \in R^n \mid Au = b,\ Du \leqslant f\}.$$

Каноническая задача линейного программирования:

$$J(u) = \langle c, u \rangle \to \inf, \quad u \in U = \{u \in R^n \mid Au = b,\ u \geqslant 0\}\,. \tag{1}$$

Любую общую задачу ЛП можно свести к канонической, правда, ценой значительного увеличения размерности. Справедлива

Теорема 1. *В канонических задачах линейного программирования 1) если допустимое множество U непусто, то оно содержит хотя бы одну угловую точку, 2) если нижняя грань функционала J_* конечна, то множество U_* оптимальных решений непусто, 3) если множество U_* оптимальных решений непусто, то оно содержит по крайней мере одну угловую точку множества U.*

Вывод: решать задачу ЛП программирования можно перебором угловых точек (вершин) канонического многогранника U. Одним из целенаправленных способов такого перебора, обеспечивающим монотонное невозрастание функционала $J(u)$, является симплекс-метод. Для его реализации нужна стартовая угловая точка (вершина) канонического многогранника U и правило распознавания угловых точек.

Алгебраический критерий для распознавания угловых точек содержится в следующей теореме.

Теорема 2. *Пусть U – канонический многогранник вида (1), A_j , $j = 1, 2, ..., n$, – столбцы матрицы A и ранг матрицы A равен $r \geqslant 1$. Тогда для того, чтобы точка $v \in U$ была угловой точкой множества U, необходимо и достаточно, чтобы равенство $Au = b$ из определения множества U выполнялось в виде*

$$A_{j_1} v_{j_1} + ... + A_{j_r} v_{j_r} = b\,, \tag{2}$$

причем столбцы A_{j_i}, $i = 1, 2, ..., r$, обязательно являются базисными для матрицы A, а не представленные в (2) координаты точки v обязательно равны нулю.

Найти стартовую угловую точку канонического многогранника U можно методом *искусственного базиса.* Для этого рассматривается следующая вспомогательная задача ЛП в пространстве переменных $z = (x, u) \in R^{m+n}$, где m – количество строк матрицы A :

$$g(z) = x_1 + x_2 + \ldots + x_m \to \inf, \quad z \in Z = \{z \geqslant 0, \;\; x + Au = b\}. \tag{3}$$

Без ограничения общности компоненты вектора b предполагаются неотрицательными: $b \geqslant 0$. Тогда по теореме 2 точка $z_0 = (x = b,\, u = 0)$ является *угловой* точкой канонического многогранника Z. Из этой точки z_0 можно запустить симплекс-метод (с антициклином) и он за конечное число шагов найдет решение $z_* = (x_*,\, v_*)$ задачи (3), которое существует в силу теоремы 1, так как $g_* = \inf\limits_{z \in Z} g(z) \geqslant 0$. При этом

$$g_* = 0 \quad \Longleftrightarrow \quad U \neq \varnothing,$$

а компонента v_* будет *угловой* точкой каконического многогранника U, из которой и запускается симплекс-метод в исходной задаче (1).

На каждом шаге симплекс-метода обрабатывается очередная угловая точка $v \in U$. Этой точке по теореме 2 соответствует *базисный* набор столбцов матрицы A с *базисными* номерами

$$J_b = \{\, j_1,\, j_2, \ldots, j_r\}.$$

Оставшиеся координаты объединяются в набор *свободных* номеров $J_f = \{1, 2, \ldots, n\} \setminus J_b$. Переменные $u_b = (u_{j_1},\, u_{j_2}, \ldots, u_{j_r})$ объявляются *базисными*, а остальные $n - r$ переменных u_f – *свободными.* Затем в задаче (1) исключаются базисные переменные и она превращается в *(неканоническую)* задачу ЛП меньшей размерности:

$$j(u_f) = J(v) - \sum_{k \in J_f} \Delta_k \, u_k \to \inf, \quad u_f \geqslant 0, \quad u_b \geqslant 0\,. \tag{4}$$

Неканоническими в задаче (4) являются ограничения $u_b \geqslant 0$ на исключенные базисные переменные. Далее, выбираются номера свободных переменных, перспективных в плане убывания функционала:

$$J_f^+ = \{\, k \in J_f \mid \; \Delta_k > 0\}.$$

При этом возможны три случая.

I. $J_f^+ = \varnothing$. Это означает, что никакой возможности уменьшить значения функционала уже нет, процесс останавливается, а искомое решение найдено: $v \in U_*, \;\; J_* = J(v)$.

II. $J_f^+ \neq \varnothing$ и $\exists k \in J_f^+$, такой, что на свободную переменную u_k неканонические условия $u_b \geqslant 0$ на самом деле никаких реальных ограничений сверху не накладывают. Тогда процесс останавливется с неутешительным выводом: решения у задачи (1) не существует, а $J_* = -\infty$.

III. В оставшемся случае, когда $J_f^+ \neq \varnothing$ и $\forall k \in J_f^+$ возможные значения переменной u_k ограничены сверху, среди этих свободных переменных выбирается одна-единственная, ей присваивается *максимально возможное* значение, а остальным свободным переменным присваиваются *нулевые* значения. В результате получается очередная *угловая* точка $w \in U$, причем $J(w) \leqslant J(v)$. Описанное здесь действие принято называть *шагом симплекс-метода*. Строгое убывание $J(w) < J(v)$ гарантировано в случае *невырожденной* угловой точки v, все базисные координаты которой *положительны*: $v_j > 0 \; \forall j \in J_b$. Если угловая точка v вырожденна, не исключено "топтание на месте": $J(w) = J(v), \;\; w = v$ и связанное с этим явлением *зацикливание*. Один из известных способов борьбы с зацикливанием симплекс-метода – *правило Блэнда*, которое упорядочивает выбор номеров свободных переменных, пригодных для одномерных вариаций в случае **III.**

Разумеется, для решения задач ЛП *с неточными данными* нужно подключать дополнительные регуляризирующие процедуры (в лекционном курсе эти проблемы не рассматривались).

Вопрос:

5. Итерационные методы минимизации: скорейшего спуска, проекции градиента и Ньютона.

План ответа.

Метод скорейшего спуска. Постановка задачи (без ограничений):

$$J(u) \to \inf, \quad u \in H, \tag{5}$$

где H – гильбертово пространство, $\dim H \leqslant \infty$, $J(u) \in C^1(H)$ – дифференцируемый функционал. Описание итерационной вычислительной процедуры метода скорейшего спуска:

$$u_{k+1} = u_k - \alpha_k J'(u_k), \quad \alpha_k = \operatorname*{argmin}_{\alpha \geqslant 0} J(u_k - \alpha J'(u_k)), \quad k = 0, 1, 2, ... \tag{6}$$

Теорема 3. *Пусть функция $J(u)$ сильно выпукла с коэффициентом $\mu > 0$ и непрерывно дифференцируема на всем гильбертовом пространстве H, а ее градиент $J'(u)$ удовлетворяет на H условию Липшица с константой $L > 0$. Тогда метод скорейшего спуска (6) сходится сильно в пространстве H из любого начального приближения $u_0 \in H$ к точке минимума u_*, причем справедлива оценка (линейной) скорости сходимости:*

$$\frac{\mu}{2}\|u_k - u_*\|^2 \leqslant J(u_k) - J(u_*) \leqslant q^k(J(u_0) - J(u_*)), \quad k = 1, 2, ...,$$

где $q = 1 - \frac{\mu}{L} \in [0, 1)$.

В случае квадратичных функционалов вида $J(u) = \|Au - f\|^2$ шаги спуска α_k в расчетных формулах (6) находятся явно.

Метод проекции градиента. Постановка задачи (с ограничениями):

$$J(u) \to \inf, \quad u \in U \subset H, \tag{7}$$

где H – гильбертово пространство, $\dim H \leqslant \infty$, U – заданное множество допустимых элементов, $J(u) \in C^1(U)$ – дифференцируемый функционал. Описание итерационной вычислительной процедуры метода проекции градиента (с постоянным шагом):

$$u_{k+1} = \mathrm{pr}_U\,(u_k - \alpha J'(u_k)), \quad k = 0, 1, 2, ... \tag{8}$$

Теорема 4. *Пусть множество U выпукло и замкнуто в гильбертовом пространстве H, функция $J(u)$ сильно выпукла на U с коэффициентом $\mu > 0$ и непрерывно дифференцируема на U, ее градиент $J'(u)$ удовлетворяет на U условию Липшица с константой $L > 0$, а постоянный шаг метода взят из диапазона $\alpha \in (0, 2\mu/(L^2))$. Тогда метод проекции градиента (8) сходится сильно в пространстве H из любого начального приближения $u_0 \in U$ к точке минимума u_*, причем справедлива оценка (линейной) скорости сходимости:*

$$\|u_k - u_*\| \leqslant \|u_0 - u_*\| \, q^k(\alpha), \quad k = 1, 2, ..., \tag{9}$$

где $q(\alpha) = \sqrt{1 - 2\mu\alpha + \alpha^2 L^2} \in (0, 1)$.

Выбор шага спуска α из указанного в условии теоремы диапазона гарантирует малость оценочного параметра $q(\alpha)$ и, тем самым, содержательность оценки (9).

Метод Ньютона. Постановка задачи (с ограничениями):

$$J(u) \to \inf, \quad u \in U \subset H, \tag{10}$$

где H – гильбертово пространство, $\dim H \leqslant \infty$, U – заданное множество допустимых элементов, $J(u) \in C^2(U)$ – дважды дифференцируемый функционал. Описание итерационной вычислительной процедуры классического метода Ньютона с шагом $\alpha_k = 1$:

$$u_{k+1} = \operatorname*{argmin}_{u \in U} \left(\langle J'(u_k), \, u - u_k \rangle + \frac{1}{2} \langle J''(u_k)(u - u_k), \, u - u_k \rangle \right), \; k = 0, 1, 2, ... \tag{11}$$

Целесообразно установить связь процедуры (11) с классическим методом Ньютона, предназначенным для решения скалярного уравнения $f(x) = 0$.

Теорема 6. *Пусть U – выпуклое замкнутое множество в гильбертовом пространстве H, имеющее непустую внутренность* $\operatorname{int} U$. *Пусть функция $J(u)$ сильно выпукла на U с коэффициентом $\mu > 0$, дважды непрерывно дифференцируема на U, а ее вторая производная $J''(u)$ удовлетворяет на U условию Липшица с константой $L > 0$. Пусть начальное приближение $u_0 \in U$ выбрано из достаточно малой окрестности точного оптимального решения u_* так, что выполнено условие*

$$q = \frac{L}{2\mu} \|u_0 - u_*\| < 1 \, . \tag{12}$$

Тогда метод Ньютона (11) сходится к u_* и справедлива оценка (квадратичной) скорости сходимости

$$\|u_k - u_*\| \le \frac{2\mu}{L} \cdot q^{2^k}, \quad k = 1, 2, ... \tag{13}$$

Метод Ньютона сходится не из любого начального приближения, а из достаточно малой окрестности искомого решения, а реализация каждого его шага, описанного в (11), требует значительных вычислительных затрат, поэтому часто метод Ньютона задействуют на завершающих этапах итерационных процедур, когда с помощью других менее трудоемких и чувствительных градиентных методов уже найдено неплохое приближенное решение, приемлемое для метода Ньютона в качестве начального приближения.