Лекция 6.
Многокритериальная оптимизация

Мельников Андрей Андреевич

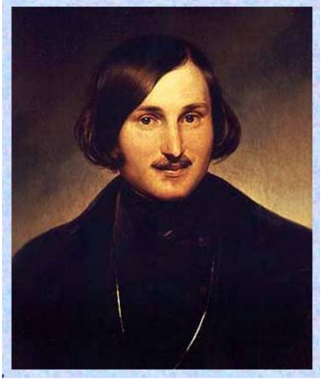
Новосибирский Государственный Университет
Факультет Информационных Технологий
http://vk.com/fit2013_tdm/

27 апреля, 2013 г.

Содержание лекции

1. Примеры
2. Основные понятия
3. Подходы к решению
4. Используемая литература

Пример



Агафья Тихоновна:
"Если бы губы Никанора Ивановича да приставить к носу Ивана Кузьмича, да взять сколько-нибудь развязности, какая у Балтазара Балтазарыча, да, пожалуй, прибавить к этому еще дородности Ивана Павловича — я бы тогда тотчас же решилась."
Н.В. Гоголь. "Женитьба". 1833

Пример "Покупка автомобиля"

	VW Golf	Opel Astra	Ford Focus	Toyota Corolla
Цена (1000 Euro)	16.2	14.9	**14.0**	15.2
Расход топлива (на 100 км)	7.2	**7.0**	7.5	8.2
Мощность (kW)	66.0	62.0	55	**71**

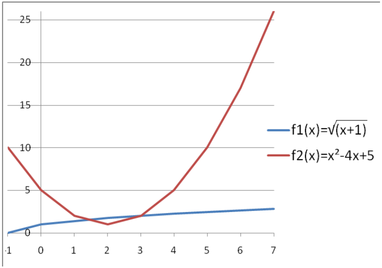
Какой автомобиль выбрать, чтобы он был мощным, недорогим, с малым расходом топлива?

Пример "Минимизация пары функций"

$$min_{x \geqslant 0} f_1(x) = \sqrt{x+1}$$

$$min_{x \geqslant 0} f_2(x) = x^2 - 4x + 5$$

минимум f_1 в точке $x_1 = 0$
минимум f_2 в точке $x_2 = 2$



Задача многокритериальной оптимизации

$$\min_{x \in X}(f_1(x), \ldots, f_p(x))$$

Основные определения

Допустимое решение $\widetilde{x} \in X$ называется слабо эффективным (слабо эффективным по Парето), если не существует решения $x \in X$ такого, что $f_k(x) < f_k(\widetilde{x})$, для всех $k = 1, ..., p$.

Множество всех слабо эффективных решений называется слабо эффективным множеством и обозначается X_{wE}.

Если $\widetilde{x}$ — слабо эффективное решение, то $\widetilde{y}$, такое что $\widetilde{y} = f(\widetilde{x})$ называется слабо недоминируемой точкой

Множество всех слабо недоминируемых точек называется слабо недоминируемым множеством и обозначается Y_{wN}.

Основные определения

Допустимое решение $\widetilde{x} \in X$ называется эффективным по Парето (оптимальным по Парето), если не существует решения $x \in X$ такого, что $f_k(x) \leqslant f_k(\widetilde{x})$, для всех $k = 1, ..., p$, и $f_k(x) < f_k(\widetilde{x})$ хотя бы для одного $i = 1, \dots, p$.

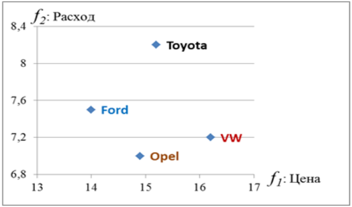
Множество всех эффективных решений называется эффективным множеством и обозначается X_E.

Если $\widetilde{x}$ — эффективное решение, то $\widetilde{y}$, такое что $\widetilde{y} = f(\widetilde{x})$ называется недоминируемой точкой.

Множество всех недоминируемых точек называется недоминируемым множеством и обозначается Y_N.

Пространство решений и пространство критериев

Пример Opel и Ford эффективный выбор или Парето*-оптимальные решения
$X = \{VW, Opel, Ford, Toyota\}$ — допустимое множество
$f_i : X \longrightarrow \mathbb{R}$, $i = 1, 2$ — критерии оптимизации $f = (f_1, f_2)$
$Y := f(X) := \{y \in \mathbb{R}^2 : y = f(x)$ для $x \in X\}$ — образ множества X или допустимое множество в пространстве критериев
$\min_{x \in X}(f_1(x), f_2(x))$



*Вильфредо Парето, 1848—1923 итальянский экономист

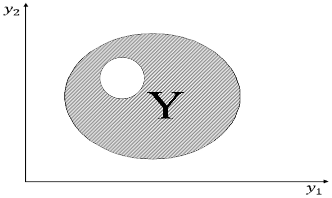
Вопрос 1

Решается задача многокритериальной оптимизации

$$f_1(x) \to \min_{x \in X}$$

$$f_2(x) \to \max_{x \in X}$$

Образ множества X в пространстве критериев приведен на рисунке.



Найдите Y_E, Y_{wE}

Линейная свертка критериев

Вместо исходной многокритериальной задачи:

$$\min_{x \in X} (f_1(x), \ldots, f_p(x))$$

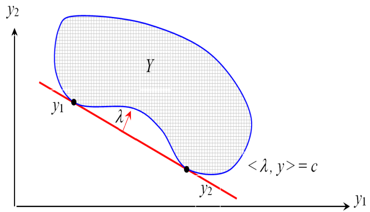
будем решать задачу с одним критерием:

$$\min_{x \in X} \sum_{k=1}^{p} \lambda_k f_k(x)$$

при разных значениях $\lambda_k \geqslant 0$ и $\sum_{k=1}^{p} \lambda_k = 1$.

Графическая интерпретация

При заданных $\lambda_k, k = 1, \ldots, p$ ищем элементы множества
$S(\lambda, Y) = \{\widetilde{y} \in Y : \langle \lambda, \widetilde{y} \rangle = \min_{y \in Y} \langle \lambda, y \rangle\}$



- Всегда ли такой процесс дает недоминируемые точки?
- Если да, то все ли точки можно получить, меняя λ_k?

Эффективность по Джеоффриону

Допустимое решение $\widetilde{x} \in X$ *называется* *эффективным по Джеоффриону* *(Arthur M. Geoffrion), если* $\widetilde{x}$ *является эффективным и существует такое положительное число* M*, что для любого* $x \in X$*, удовлетворяющего условию*

$$f_i(x) < f_i(\widetilde{x}) \text{ для некоторого } i,$$

найдется такой индекс j*, что* $f_j(\widetilde{x}) < f_j(x)$ *и выполнено неравенство*

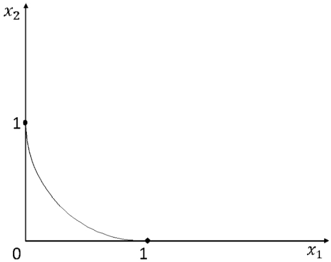
$$\frac{f_i(\widetilde{x}) - f_i(x)}{f_j(x) - f_j(\widetilde{x})} \leqslant M.$$

Точка $\widetilde{y}$*, такая что* $\widetilde{y} = f(\widetilde{x})$ *называется* *недоминируемой по Джеоффриону*.

Вопрос 2.

Пусть $X = \{(x_1, x_2) \in \mathbb{R}^2 : (x_1 - 1)^2 + (x_2 - 1)^2 = 1, x_1, x_2 \in [0, 1]$.
Для $x = (x_1, x_2) \in X$ положим

$$f_1(x) = x_1, \quad f_2(x) = x_2$$



Какие точки являются эффективными по Джеоффриону?

Свойства линейной свертки

Теорема 1

Пусть $\lambda_k > 0, k = 1, \ldots, p$ *и* $\sum_{k=1}^{p} \lambda_k = 1$. *Если* $\widetilde{x}$ *— оптимальное решение линейной свертки с коэффициентами* λ_k, *то* $\widetilde{x}$ *— эффективное по Джеоффриону решение.*

Свойства линейной свертки

Теорема 1

Пусть $\lambda_k > 0, k = 1, \ldots, p$ и $\sum_{k=1}^{p} \lambda_k = 1$. Если $\widetilde{x}$ — оптимальное решение линейной свертки с коэффициентами λ_k, то $\widetilde{x}$ — эффективное по Джеоффриону решение.

Доказательство.
Покажем, что $\widetilde{x}$ — эффективное решение. Если это не так, то найдется $x' \in X : f_i(x') \leqslant f_i(\widetilde{x})$ и $\exists i_0 : f_{i_0}(x') < f_{i_0}(\widetilde{x})$.
В таком случае

$$\sum_{k=1}^{p} \lambda_k f_k(x') < \sum_{k=1}^{p} \lambda_k f_k(\widetilde{x}),$$

что противоречит оптимальности $\widetilde{x}$ в линейной свертке.

Покажем эффективность по Джеоффриону.
Положим $M = (p-1)\max\limits_{i,j}\frac{\lambda_j}{\lambda_i}$.
Пусть найдутся $x \in X$ и критерий f_i такие, что $f_i(x) < f_i(\widetilde{x})$ и

$$f_i(\widetilde{x}) - f_i(x) > M(f_j(x) - f_j(\widetilde{x}))$$

для всех критериев f_j, для которых $f_j(\widetilde{x}) < f_j(x)$.
Тогда неравенство

$$f_i(\widetilde{x}) - f_i(x) > (p-1)\frac{\lambda_j}{\lambda_i}(f_j(x) - f_j(\widetilde{x})).$$

верно для всех $j \neq i$, т.к. при $f_j(\widetilde{x}) \geqslant f_j(x)$ оно тривиально.
Умножим это неравенство на $\frac{\lambda_i}{(p-1)}$ и просуммируем по всем $j \neq i$:

$$\lambda_i(f_i(\widetilde{x}) - f_i(x)) > \sum_{j\neq i}\lambda_j(f_j(x) - f_j(\widetilde{x})),$$

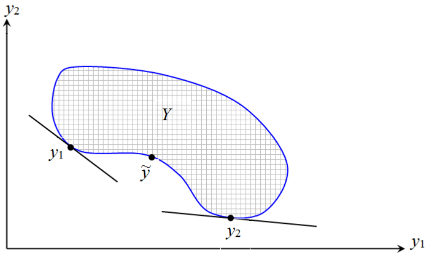
что противоречит оптимальности $\widetilde{x}$.

Свойства линейной свертки.

Верно ли обратное утверждение?

Свойства линейной свертки.

Верно ли обратное утверждение?



$\tilde{y}$ — недоминируемая точка

Вспомогательная лемма

Лемма (о свойствах выпуклых функций)

Пусть $X \subset \mathbb{R}^n$ *— выпуклое множество и все функции* $h_k : \mathbb{R}^n \to \mathbb{R}$ *выпуклые,* $k = 1, \dots, p$.
Если система $h_k(x) < 0$, $k = 1, \dots, p$, *на* X *не имеет решений, то* $\exists \{\lambda_k\} : \lambda_k > 0, k = 1, \dots, p; \sum_{k=1}^{p} \lambda_k = 1$ *т.ч.* $\sum_{k=1}^{p} \lambda_k h_k(x) \geqslant 0$ *для всех* $x \in X$.

Без доказательства.

Свойства линейной свертки.

Теорема 2

Пусть $X \subset \mathbb{R}^n$ — выпуклая область и все функции $f_k : \mathbb{R}^n \to \mathbb{R}$ выпуклые, $k = 1, \ldots, p$. Тогда $\widetilde{x} \in X$ является эффективным решением по Джеоффриону, если и только если $\widetilde{x}$ — оптимальное решение линейной свертки с весами $\lambda_k > 0$, $k = 1, \ldots, p$.

Свойства линейной свертки.

Теорема 2

Пусть $X \subset \mathbb{R}^n$ — выпуклая область и все функции $f_k : \mathbb{R}^n \to \mathbb{R}$ выпуклые, $k = 1, \ldots, p$. Тогда $\widetilde{x} \in X$ является эффективным решением по Джеоффриону, если и только если $\widetilde{x}$ — оптимальное решение линейной свертки с весами $\lambda_k > 0$, $k = 1, \ldots, p$.

Доказательство.
Проверим необходимость. Достаточность следует из теоремы 1.
Пусть $\widetilde{x}$ — эффективное решение по Джеоффриону.
Из определения следует, что существует $M > 0$, для которого при любом $i = 1, \ldots, p$ система

$$f_i(x) < f_i(\widetilde{x})$$

$$f_i(x) + M f_j(x) < f_i(\widetilde{x}) + M f_j(\widetilde{x}), j \neq i$$

не имеет решений.

Тогда по лемме о выпуклых функциях для i-й системы найдутся величины $\lambda_k^i > 0, k = 1, \ldots, p$ т.ч. $\sum_{k=1}^{p} \lambda_k^i = 1$, при которых для любого $x \in X$ верно неравенство:

$$\lambda_i^i f_i(x) + \sum_{k \neq i} \lambda_k^i (f_i(x) + M f_k(x)) \geqslant \lambda_i^i f_i(\widetilde{x}) + \sum_{k \neq i} \lambda_k^i (f_i(\widetilde{x}) + M f_k(\widetilde{x})).$$

Тогда по лемме о выпуклых функциях для i-й системы найдутся величины $\lambda_k^i > 0, k = 1, \ldots, p$ т.ч. $\sum_{k=1}^p \lambda_k^i = 1$, при которых для любого $x \in X$ верно неравенство:

$$\lambda_i^i f_i(x) + \sum_{k \neq i} \lambda_k^i (f_i(x) + M f_k(x)) \geqslant \lambda_i^i f_i(\widetilde{x}) + \sum_{k \neq i} \lambda_k^i (f_i(\widetilde{x}) + M f_k(\widetilde{x})).$$

$$f_i(x) + M \sum_{k \neq i} \lambda_k^i f_k(x) \geqslant f_i(\widetilde{x}) + M \sum_{k \neq i} \lambda_k^i f_k(\widetilde{x}), i = 1, \ldots, p.$$

Тогда по лемме о выпуклых функциях для i-й системы найдутся величины $\lambda_k^i > 0, k = 1, \ldots, p$ т.ч. $\sum_{k=1}^p \lambda_k^i = 1$, при которых для любого $x \in X$ верно неравенство:

$$\lambda_i^i f_i(x) + \sum_{k \neq i} \lambda_k^i (f_i(x) + M f_k(x)) \geqslant \lambda_i^i f_i(\widetilde{x}) + \sum_{k \neq i} \lambda_k^i (f_i(\widetilde{x}) + M f_k(\widetilde{x})).$$

$$f_i(x) + M \sum_{k \neq i} \lambda_k^i f_k(x) \geqslant f_i(\widetilde{x}) + M \sum_{k \neq i} \lambda_k^i f_k(\widetilde{x}), i = 1, \ldots, p.$$

$$\sum_{i=1}^p f_i(x) + M \sum_{i=1}^p \sum_{k \neq i} \lambda_k^i f_i(x) \geqslant \sum_{i=1}^p f_i(\widetilde{x}) + M \sum_{i=1}^p \sum_{k \neq i} \lambda_k^i f_i(\widetilde{x}).$$

Тогда по лемме о выпуклых функциях для i-й системы найдутся величины $\lambda_k^i > 0, k = 1, \dots, p$ т.ч. $\sum_{k=1}^p \lambda_k^i = 1$, при которых для любого $x \in X$ верно неравенство:

$$\lambda_i^i f_i(x) + \sum_{k \neq i} \lambda_k^i (f_i(x) + M f_k(x)) \geqslant \lambda_i^i f_i(\widetilde{x}) + \sum_{k \neq i} \lambda_k^i (f_i(\widetilde{x}) + M f_k(\widetilde{x})).$$

$$f_i(x) + M \sum_{k \neq i} \lambda_k^i f_k(x) \geqslant f_i(\widetilde{x}) + M \sum_{k \neq i} \lambda_k^i f_k(\widetilde{x}), i = 1, \dots, p.$$

$$\sum_{i=1}^p f_i(x) + M \sum_{i=1}^p \sum_{k \neq i} \lambda_k^i f_i(x) \geqslant \sum_{i=1}^p f_i(\widetilde{x}) + M \sum_{i=1}^p \sum_{k \neq i} \lambda_k^i f_i(\widetilde{x}).$$

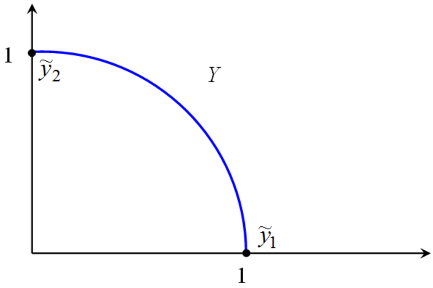
$$\sum_{k=1}^p (1 + M \sum_{k \neq i} \lambda_k^i) f_k(x) \geqslant \sum_{k=1}^p (1 + M \sum_{k \neq i} \lambda_k^i) f_k(\widetilde{x})$$

верно для всех $x \in X$. Поделив обе части на $\sum_{k=1}^p (1 + M \sum_{k \neq i} \lambda_k^i)$, получаем нормированный вектор $\lambda > 0$, указанный в теореме, при котором $\widetilde{x}$ — оптимальное решение в линейной свертке.

Пример

$$X = \{x \in \mathbb{R}^2_{\geq} : x_1^2 + x_2^2 \geqslant 1\}$$

$$f_1(x) = x_1, f_2(x) = x_2$$



$$X_E = \{x \in X : x_1^2 + x_2^2 = 1\}$$

При любых $\lambda \geqslant 0$ линейная свертка дает только $\widetilde{y}_1, \widetilde{y}_2$!

Метод уступок

Для $b \in \mathbb{R}^p$ рассмотрим задачу $*$ с одним критерием:

$$\min_{x \in X} f_j(x)$$

при условии

$$f_k(x) \leqslant b, k \neq j.$$

Метод уступок

Теорема 3

Допустимое решение $\widetilde{x} \in X$ *является эффективным по Парето* $\Leftrightarrow$ $\exists b \in \mathbb{R}^p$, *при котором* $\widetilde{x}$ *— оптимальное решение задачи* $*$ *для всех* $j = 1, \ldots, p$.

Метод уступок

Теорема 3

Допустимое решение $\widetilde{x} \in X$ *является эффективным по Парето* $\Leftrightarrow$ $\exists b \in \mathbb{R}^p$, *при котором* $\widetilde{x}$ *— оптимальное решение задачи* $*$ *для всех* $j = 1, \ldots, p$.

Доказательство.
Пусть $\widetilde{x} \in X_E$. Положим $\widetilde{b} = f(\widetilde{x})$ и предположим, что $\widetilde{x}$ не является оптимальным решением для некоторого j. Тогда найдется $x \in X$, для которого $f_j(x) < f_j(\widetilde{x})$ и $f_k(x) \leqslant \widetilde{b}_k = f_k(\widetilde{x})$, $k \neq j$,, что противоречит нашему предположению.
Пусть $\widetilde{x} \notin X_E$. Тогда $\exists j_0$ и решение $x \in X$, для которых $f_{j_0}(x) < f_{j_0}(\widetilde{x})$ и $f_k(x) \leqslant f_k(\widetilde{x})$, $k \neq j_0$. Поэтому $\widetilde{x}$ не может быть оптимальным решением для всех j ни при каком b.

Используемая литература

Ehrgott M. Multicriteria Optimization. 2^{nd} Edition. Berlin: Springer, 2005.

Эффективность по Джеоффриону

Точка $\widehat{x} = (1, 0)$ не является эффективной по Джеоффриону

- Покажем, что для любого $M > 0$ $\exists i \in \{1, 2\}$, т.ч. для некоторого $x \in X$ с $f_i(x) < f_i(\widehat{x})$, $\frac{f_i(\widehat{x}) - f_i(x)}{f_j(x) - f_j(\widehat{x})} > M$ для любого $j \in \{1, 2\}$ с $f_j(\widehat{x}) < f_j(x)$.

Пусть $i = 1$, возьмем $(x_1^\varepsilon, x_2^\varepsilon) = (1 - \varepsilon, 1 - \sqrt{1 - \varepsilon^2})$.
$(x_1^\varepsilon, x_2^\varepsilon)$ — эффективное решение, т.к. $(x_1^\varepsilon - 1)^2 + (x_2^\varepsilon - 1)^2 = 1$.
$(x_1^\varepsilon, x_2^\varepsilon) \in X$, $x_1^\varepsilon < \widehat{x}_1$, $x_2^\varepsilon > \widehat{x}_2$, $i = 1$, $j = 2$.

$$\frac{f_i(\widehat{x}) - f_i(x^\varepsilon)}{f_j(x^\varepsilon) - f_j(\widehat{x})} = \frac{1 - (1 - \varepsilon)}{1 - \sqrt{1 - \varepsilon^2}} = \frac{\varepsilon}{1 - \sqrt{1 - \varepsilon^2}} \longrightarrow^{\varepsilon \to 0} \infty$$

Аналогично для $\bar{x} = (0, 1)$.