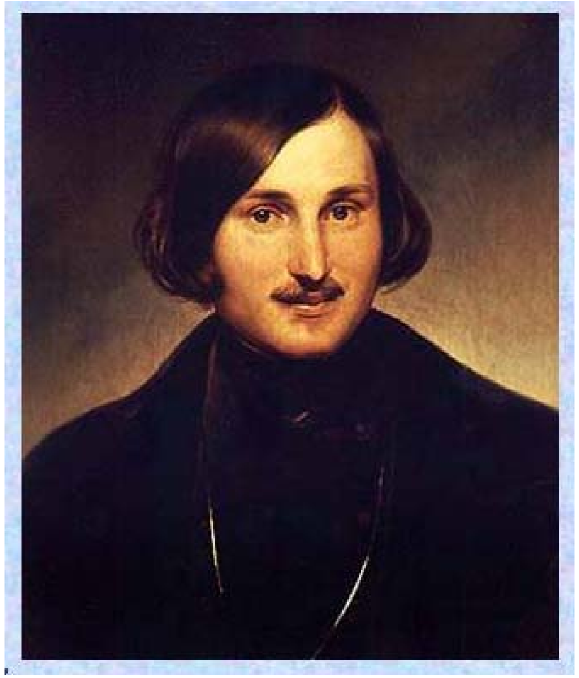
Многокритериальная оптимизация



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

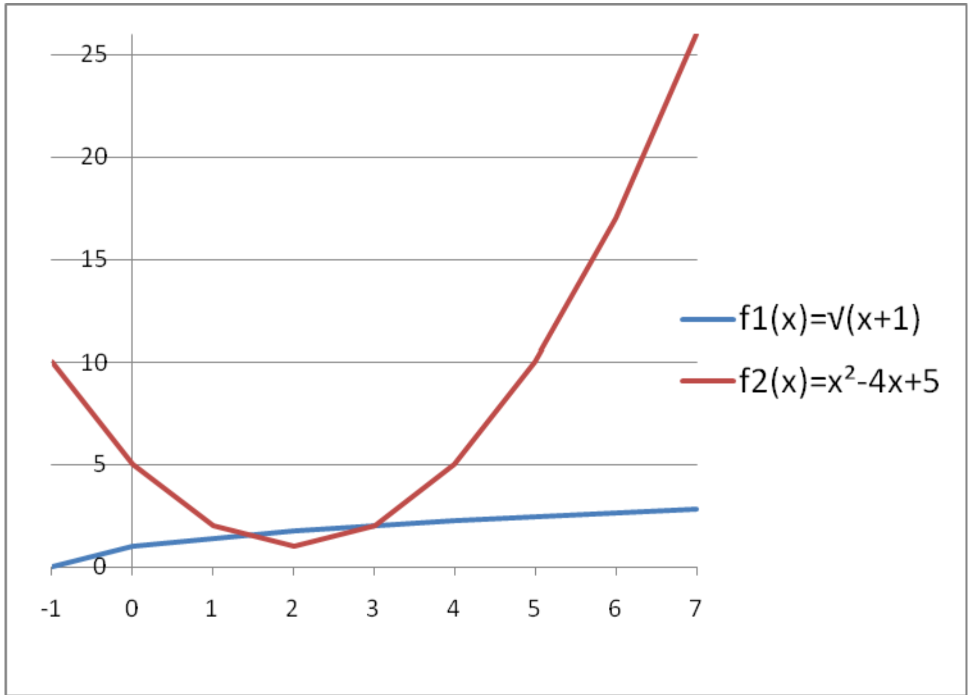
минимизация пары функций

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

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

минимум f_1 в точке $x_1 = 0$

минимум f_2 в точке $x_2 = 2$



Задача многокритериальной оптимизации

$$\min(f_1(x),\dots,f_p(x))$$

при условии $\qquad x \in X.$

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

Если $\tilde{x}$ — эффективное решение, то $\tilde{y} = f(\tilde{x})$ называется *недоминируемой точкой*. Множество всех недоминируемых точек называется *недоминируемым множеством* и обозначается Y_N. Множество всех эффективных решений называется *эффективным множеством* и обозначается X_E.

Пространство решений и пространство критериев

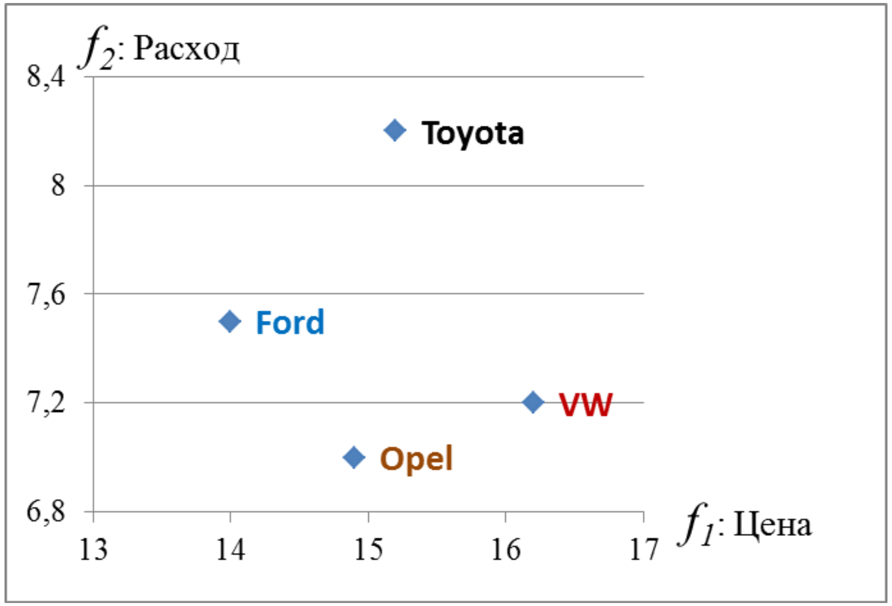
Пример **Opel** и **Ford** эффективный выбор или Парето*-оптимальные решения

$$X = \{VW, Opel, Ford, Toyota\}$$

допустимое множество

$$f_i : X \to \mathrm{R}, i = 1, 2$$

критерии оптимизации $f = (f_1, f_2)$



$$Y := f(X) := \{y \in \mathrm{R}^2 : y = f(x) \text{ для } x \in X\}$$

образ множества X или допустимое множество в пространстве критериев

$$\min_{x \in X} (f_1(x), f_2(x))$$

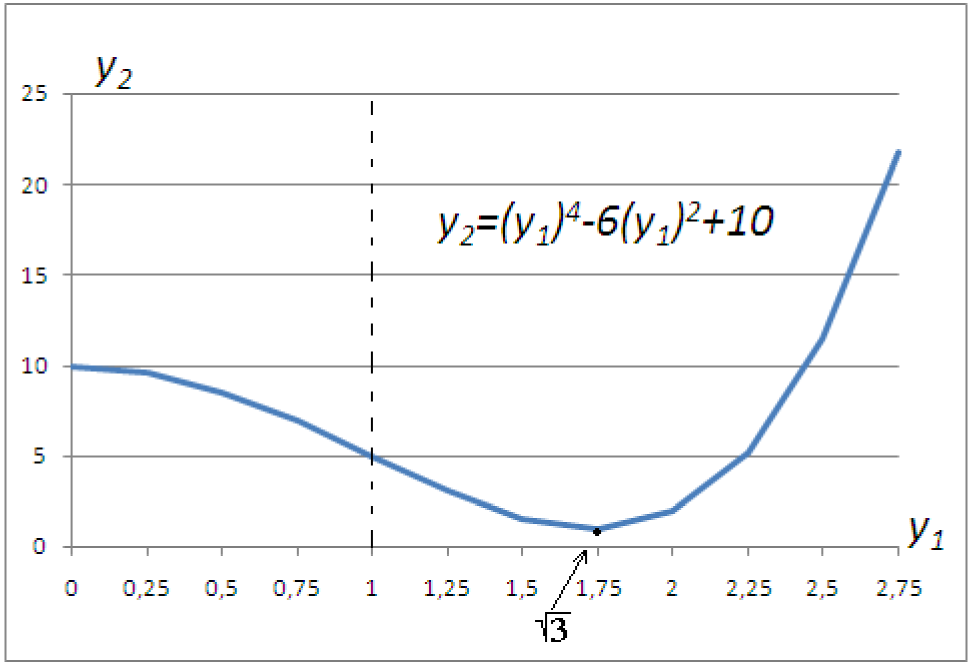
*Вильфредо Парето, 1848-1923 итальянский экономист

Пространство решений и пространство критериев

Пример минимизации двух функций

допустимое множество
$X = \{x \in \mathrm{R} : x \geq 0\}$

допустимое множество в пространстве критериев Y состоит из части графика, расположенной справа от вертикальной линии $y_1 = 1$



недоминируемые точки: $1 \leq y_1 \leq \sqrt{3}$, $1 \leq y_2 \leq 5$, соответствующие решениям $x \in [0,2]$

Линейная свертка критериев

Вместо исходной многокритериальной задачи

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

будем решать задачу с одним взвешенным критерием

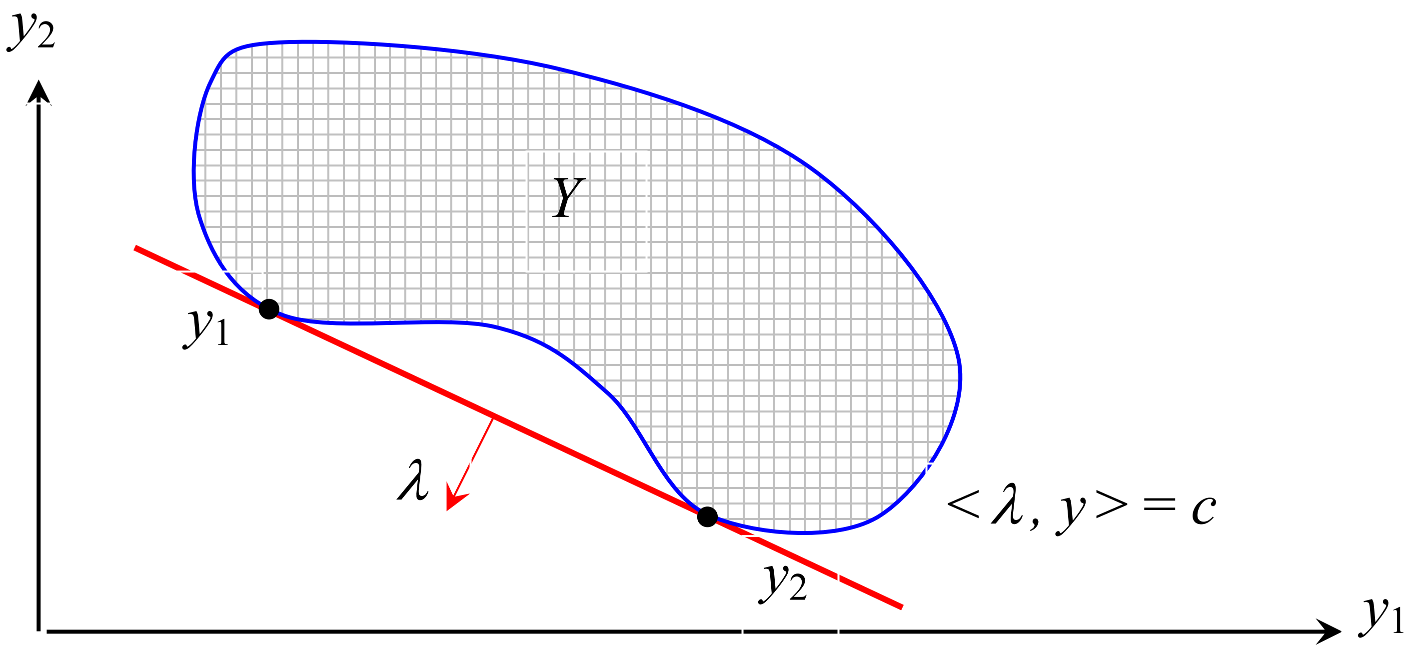
$$\min_{x \in X} \sum_{k=1}^{p} \lambda_k f_k(x)$$

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

Графическая интерпретация

При заданных λ ищем элементы множества

$$S(\lambda, Y) = \{\tilde{y} \in Y : <\lambda, \tilde{y}> = \min_{y \in Y} <\lambda, y>\}$$



1. Всегда ли такой процесс дает недоминируемые точки?

2. Если да, то все ли точки можно получить, меняя λ ?

Эффективность по Джеоффриону

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

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

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

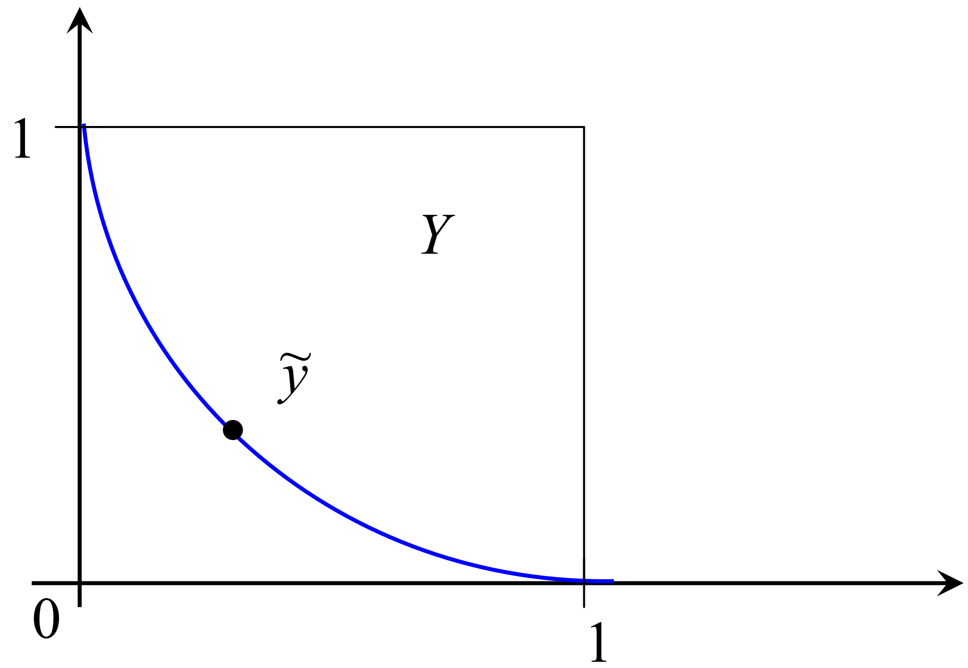
$$\frac{f_i(\tilde{x}) - f_i(x)}{f_j(x) - f_j(\tilde{x})} \le M .$$

Точка $\tilde{y} = f(\tilde{x})$ называется *недоминируемой по Джеоффриону.*

Пример.

$$X = \{(x_1, x_2) \in R^2 : (x_1 - 1)^2 + (x_2 - 1)^2 \le 1,\quad 0 \le x_1, x_2 \le 1\}$$

$Y = X$



$\tilde{y}$ — недоминируемая точка

$x = (0,\ 1)$ не является эффективным по Джеоффриону.

Теорема. Пусть положительные величины λ_k, $k = 1,\ldots, p$ удовлетворяют равенству $\sum_{k=1}^{p} \lambda_k = 1$. Если $\tilde{x}$ — оптимальное решение линейной свертки, то $\tilde{x}$ — эффективное решение по Джеоффриону.

Доказательство. Покажем, что $\tilde{x}$ — эффективное решение. Пусть $x' \in X$, $f(x') \le f(\tilde{x})$ и существует индекс i такой, что $f_i(x') < f_i(\tilde{x})$. Т.к. $\lambda_k > 0$, то

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

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

Покажем эффективность по Джеоффриону. Положим $M := (p-1)\max\limits_{ij} \frac{\lambda_j}{\lambda_i}$.

Предположим, существует $x \in X$ и такой индекс $i \le p$, что $f_i(x) < f_i(\tilde{x})$ и $f_i(\tilde{x}) - f_i(x) > M(f_j(x) - f_j(\tilde{x}))$ для всех индексов j, где $f_j(\tilde{x}) < f_j(x)$.

Тогда по выбору M получаем

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

Заметим, что неравенство верно для всех $j \neq i$, т.к. при $f_j(\tilde{x}) \geq f_j(x)$ оно тривиально. Умножим это неравенство на $\lambda_i / (p-1)$ и сложим по всем $j \neq i$:

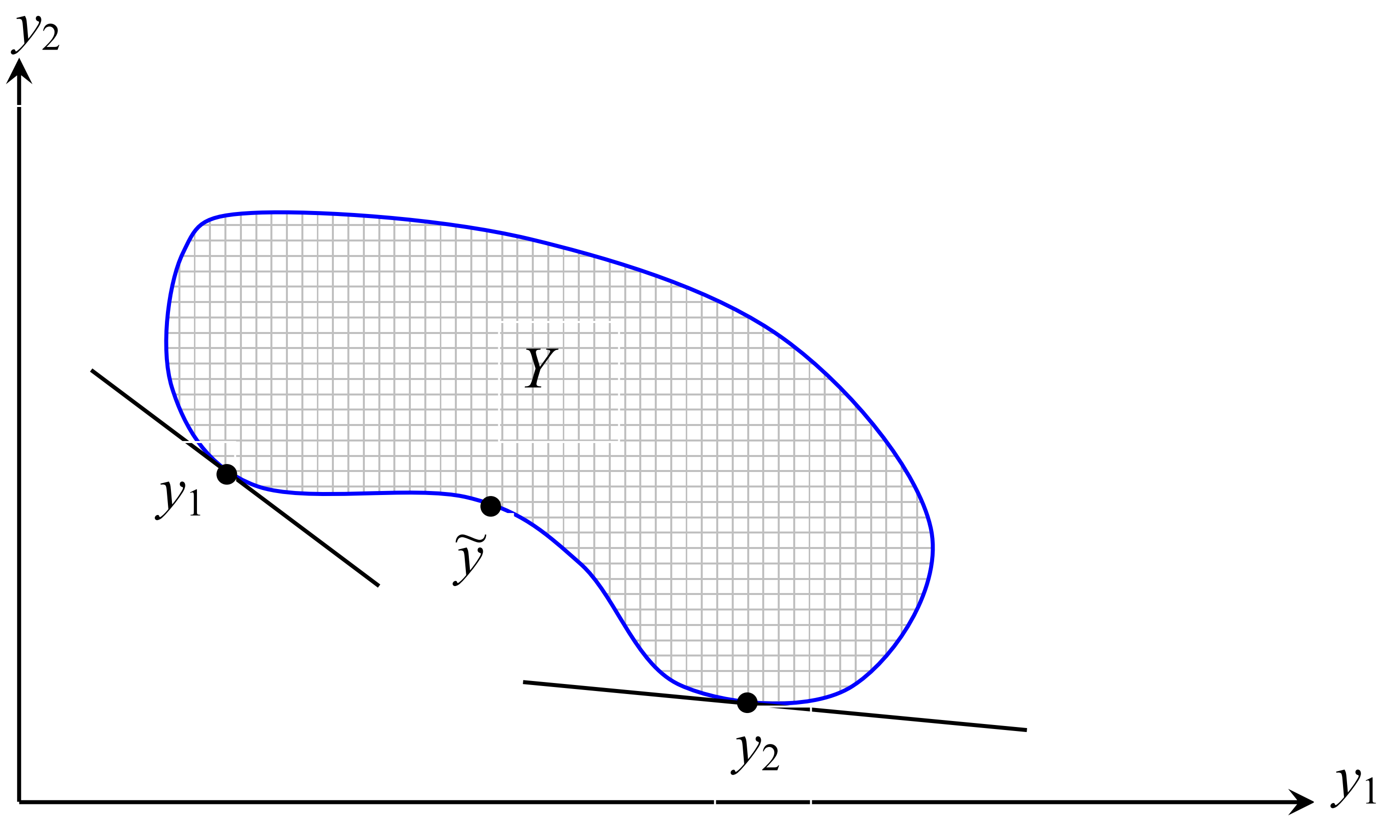
$$\lambda_i (f_i(\tilde{x}) - f_i(x)) > \sum_{j \neq i} \lambda_j (f_j(x) - f_j(\tilde{x}))$$

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

группируем $\lambda_i f_i(\tilde{x}) + \sum_{j \neq i} \lambda_j f_j(\tilde{x}) > \lambda_i f_i(x) + \sum_{j \neq i} \lambda_j f_j(x)$.

Получаем $\sum_{i=1}^{p} \lambda_i f_i(\tilde{x}) > \sum_{i=1}^{p} \lambda_i f_i(x)$, что противоречит оптимальности $\tilde{x}$. ■

Верно ли обратное утверждение?



$\tilde{y}$ — недоминируемая точка

Лемма (о свойствах выпуклых функций). Пусть $X \subset R^n$ — выпуклое множество и все функции $h_k : R^n \to R$ выпуклые $k = 1,\ldots, p$. Если система $h_k(x) < 0$, $k = 1,\ldots, p$, не имеет решений x из множества X, то существуют такие неотрицательные величины λ_k, в сумме равные 1, $\left(\sum_{k=1}^{p} \lambda_k = 1,\ \lambda_k \geq 0 \right)$, что

$$\sum_{k=1}^{p} \lambda_k h_k(x) \geq 0 \quad \text{для всех } x \in X.$$

Без доказательства.

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

Доказательство. Проверим необходимость. Достаточность следует из предыдущей теоремы.

Пусть $\tilde{x}$ — эффективное решение по Джеоффриону. Из определения следует, что существует $M > 0$, для которого при любом $i = 1,\ldots, p$ система

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

$$f_i(x) + Mf_j(x) < f_i(\tilde{x}) + Mf_j(\tilde{x}), \quad j \neq i$$

не имеет решений.

Тогда по лемме о выпуклых функциях для i-й системы найдутся величины $\lambda_k^i \geq 0$, $k = 1,\ldots, p$ в сумме равные 1, т.е. $\sum\limits_{k=1}^{p} \lambda_k^i = 1$ при которых для любого $x \in X$ верно неравенство:

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

Открываем скобки:

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

Заносим в сумму первое слагаемое в обеих частях неравенства:

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

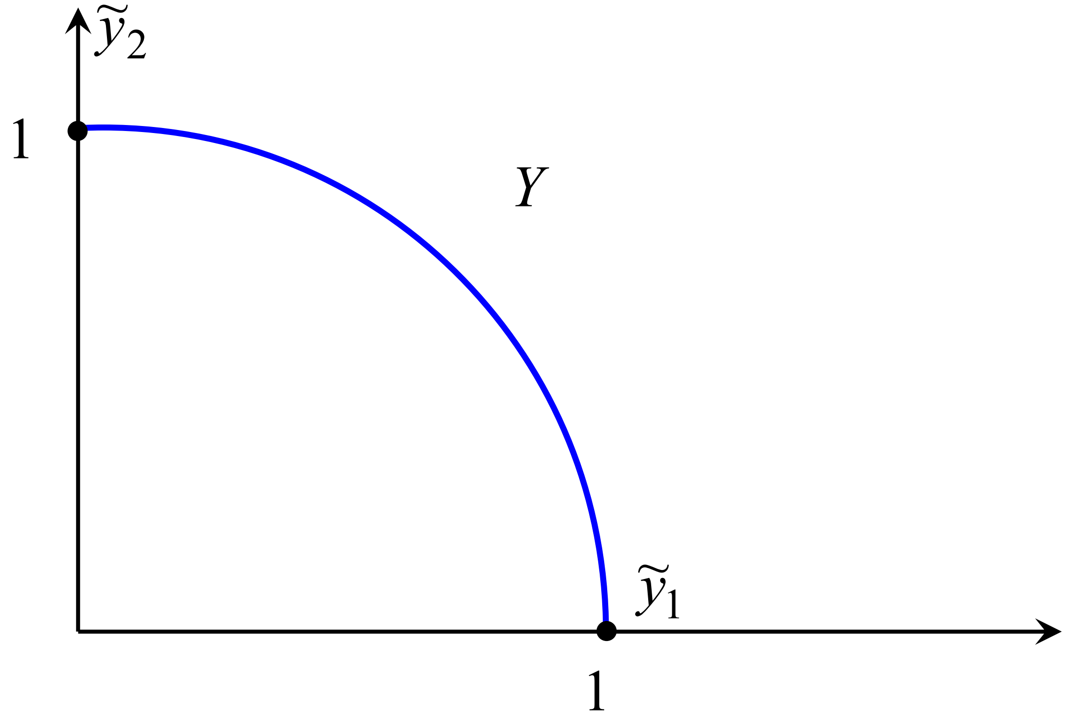
Пользуясь равенством $\sum_{k=1}^{p} \lambda_k^i = 1$, получаем:

$$f_i(x) + M\sum_{k \neq i} \lambda_k^i f_k(x) \geq f_i(\tilde{x}) + M\sum_{k \neq i} \lambda_k^i f_k(\tilde{x}).$$

Итак, для каждого $i = 1,\ldots, p$ получили неравенство. Складывая их по i, получаем: $\sum_{i=1}^{p} f_i(x) + M\sum_{i=1}^{p}\sum_{k \neq i} \lambda_k^i f_i(x) \geq \sum_{i=1}^{p} f_i(\tilde{x}) + M\sum_{i=1}^{p}\sum_{k \neq i} \lambda_k^i f_i(\tilde{x}).$

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

Пример. $X = \{x \in R^2_{\geq} \mid x_1^2 + x_2^2 \geq 1\} \quad f_1(x) = x_1, f_2(x) = x_2$



$X_E = \{x \in X \mid x_1^2 + x_2^2 = 1\}$ — решения, эффективные по Парето.

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

Метод уступок

Для $\varepsilon \in R^p$ рассмотрим задачу с одним критерием

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

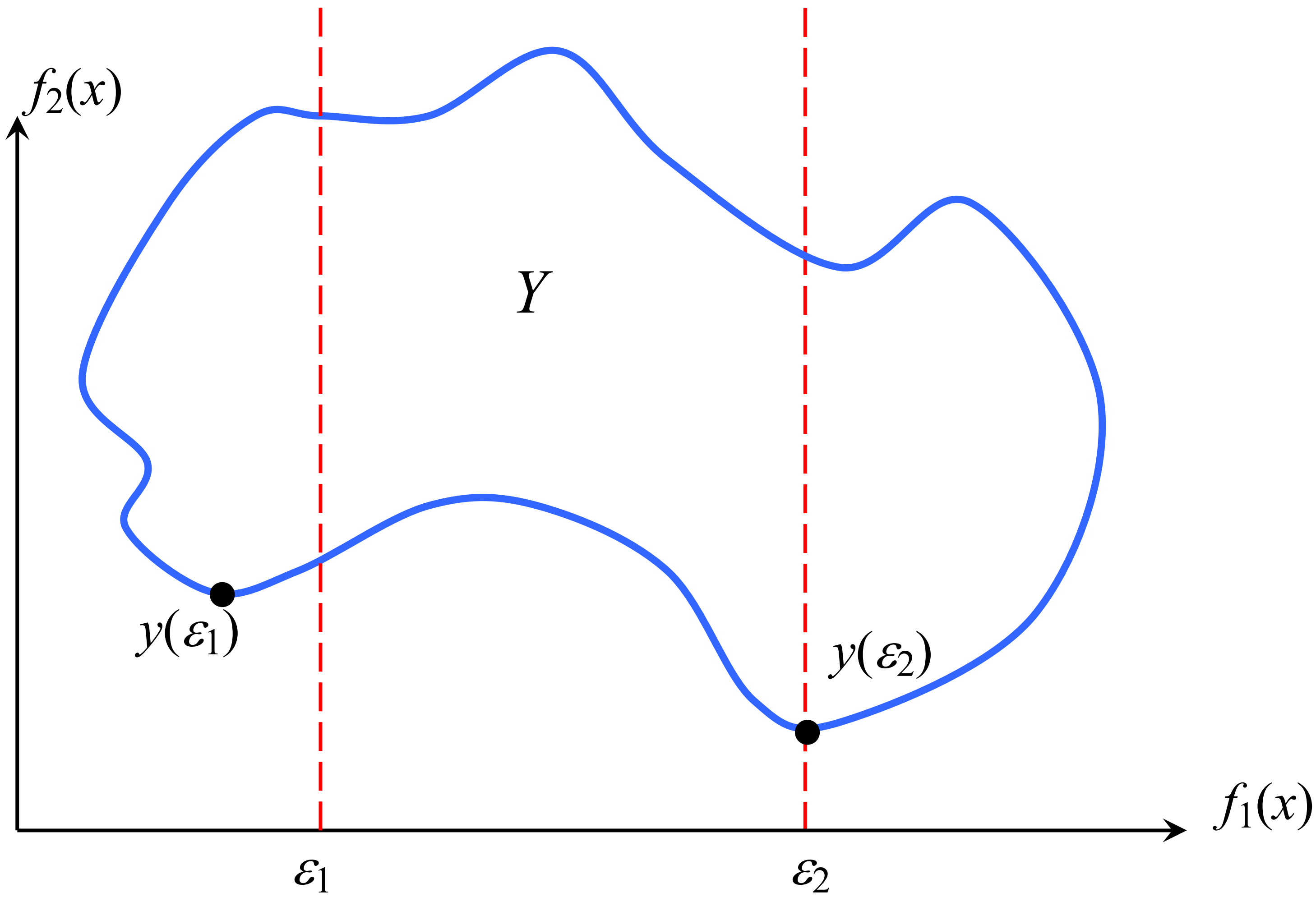
при условии $\quad f_k(x) \le \varepsilon, \;\; k \ne j \qquad (*)$

Теорема. Допустимое решение $\tilde{x} \in X$ является эффективным по Парето $\Leftrightarrow$ $\exists\; \varepsilon \in R^p$, при котором $\tilde{x}$ — оптимальное решение (*) для всех $j = 1,\dots, p$.

Доказательство. $\Rightarrow$ Положим $\tilde{\varepsilon} = f(\tilde{x})$ и предположим, что $\tilde{x}$ не является оптимальным решением для некоторого j. Тогда найдется $x \in X$, для которого $f_j(x) < f_j(\tilde{x})$ и $f_k(x) \le \tilde{\varepsilon}_k = f_k(\tilde{x})$, $k \ne j$, т.е. $\tilde{x}$ не является эффективным по Парето.

$\Leftarrow$ Предположим, что $\tilde{x} \notin X_E$. Тогда $\exists\; j$ и решение $x \in X$, для которых $f_j(x) < f_j(\tilde{x})$ и $f_k(x) \le f_k(\tilde{x})$, $k \ne j$. Поэтому $\tilde{x}$ не может быть оптимальным решением ни при каком ε, если $\tilde{x}$ — допустимое решение для этого ε. ■

Пример



$f_2(x)$
Y
$y(\varepsilon_1)$
$y(\varepsilon_2)$
$f_1(x)$
ε_1
ε_2