#9. Непрерывные оптимизационные модели

Основные концепции однокритериальных и многокритериальных задач

Постановка задачи оптимизации

- Необходимо найти минимум (в общем случае экстремум) функции f:
$$\min_{X \in D} f(\mathbf{X}) = f(\mathbf{X}^*) = f^*.$$
- $\mathbf{X} \in \mathbb{R}^n$ – вектор параметров:
$$\mathbf{X} = (x_1, x_1, \dots, x_n)^{\mathrm{T}},$$
здесь $n = |\mathbf{X}|$ - размер вектора **X**.
- $f(\mathbf{X}) \rightarrow \mathbb{R}$ – целевая функция (критерий)
- D – область допустимых значений вектора **X**, формируемая ограничениями типа равенств и неравенств:
$$D = \{X \mid E(\mathbf{X}) = 0, G(\mathbf{X}) \geq 0\}.$$

Особенности целевых функций (ЦФ)

- ***Унимодальная*** ЦФ – имеет только один экстремум на рассматриваемом отрезке (для n = 1).
- ЦФ называется ***выпуклой*** (вниз), если для любых векторов $\mathbf{X}_1, \mathbf{X}_2 \in D$, $\mathbf{X}_1 \neq \mathbf{X}_2$ и любого $\lambda \in [0,1]$:
$$f(\lambda\mathbf{X}_1 + (1-\lambda)\mathbf{X}_2) \leq f(\lambda\mathbf{X}_1) + (1-\lambda)f(\mathbf{X}_2).$$
- ЦФ, имеющую в своей области определения несколько локальных минимумов, называют ***многоэкстремальной***.
- Целевая функция называется ***сепарабельной***, если
$$f(\mathbf{X}) = \sum_{i=1}^{n} f_i(x_i)\,.$$
- ЦФ называется ***позиномиальной***, если
$$f(\mathbf{X}) = \sum_{i=1}^{k} c_i p_i(\mathbf{X})\,, \quad p_i(X) = \prod_{j=1}^{n} x_j^{a_{ij}}\,.$$

Классификация задач оптимизации

- По виду целевой функции и ограничивающих функций:
  - Задача линейного / дробно-линейного программирования
  - Задача сепарабельного / геометрического программирования
  - Задача выпуклого / квадратичного программирования
  - Задача дискретного / целочисленного программирования
  - Задача нелинейного программирования
- По наличию или отсутствию ограничений:
  - Задача безусловной / условной оптимизации
- По характеру ограничений:
  - Задача с ограничениями типа равенств / неравенств
  - Задача с ограничениями общего вида
- По размерности вектора **X**:
  - Задача однопараметрической / многопараметрической оптимизации
- По числу точек экстремума:
  - Задача одноэкстремальная / многоэкстремальная
- По характеру искомого решения:
  - Задача локальной / глобальной оптимизации

Алгоритмы оптимизации

- ***Испытанием*** называют операцию однократного вычисления ЦФ $f(\mathbf{X})$ и ограничивающих функций $E(\mathbf{X})$, $G(\mathbf{X})$, а также, быть может, производных указанных функций в некоторой точке $X \in D$.
- Особенность большинства задач оптимизации в том, что каждое испытание может потребовать больших затрат машинного времени.
- Приближение к решению задачи (кандидата на решение) назовем ***агентом***.
- В большинстве случаев используют ***итерационные алгоритмы оптимизации***, в которых новое положение агентов вычисляются с использованием миграционного оператора $\Phi(\cdot)$:

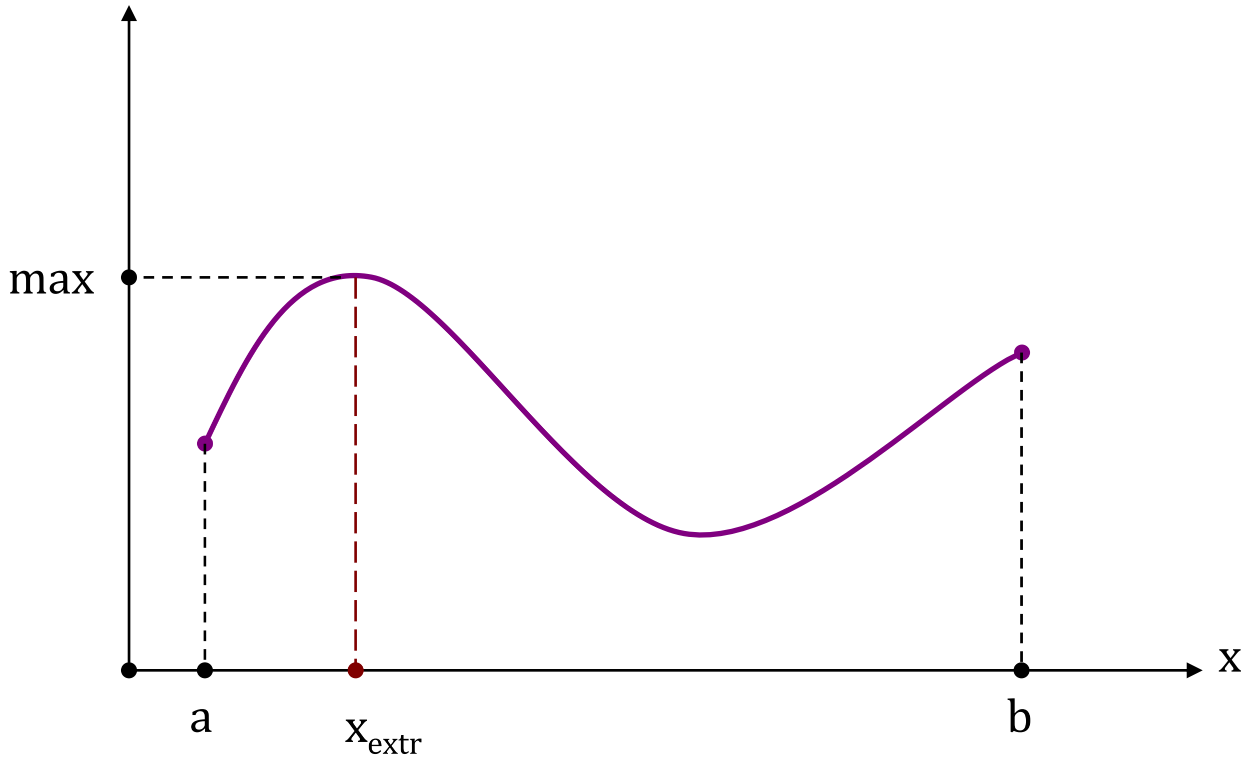
$$X_i^{t+1} = \Phi\left(X_i^t, f\left(X_i^t\right), E\left(X_i^t\right), G\left(X_i^t\right)\right).$$

Классификация алгоритмов оптимизации

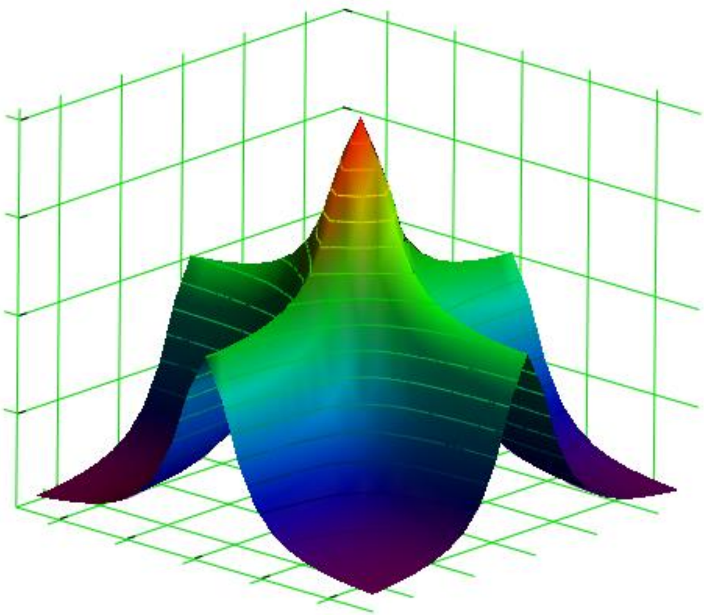
- По характеру искомого решения:
  - Алгоритм локальной / глобальной оптимизации
- По характеру ограничений:
  - Алгоритм безусловной / условной оптимизации
- По характеру миграционного оператора $\Phi(\cdot)$:
  - Детерминированный / стохастический алгоритм оптимизации
- По классу алгоритма:
  - Пассивный / последовательный алгоритм оптимизации
- По числу предыдущих учитываемых шагов:
  - Одношаговый / многошаговый алгоритм
- По порядку используемых производных:
  - Алгоритм нулевого / k-го порядка
- По числу агентов:
  - Траекторные / популяционные алгоритмы

Экстремум функции одной переменной (n = 1)

- Условие экстремума: $f(x) \to \max$
- Необходимое условие: $f'(x) = 0$.
- Достаточное условие (max): $f''(x) < 0$ $f'(x) \leq 0$.



Экстремум функции нескольких переменных

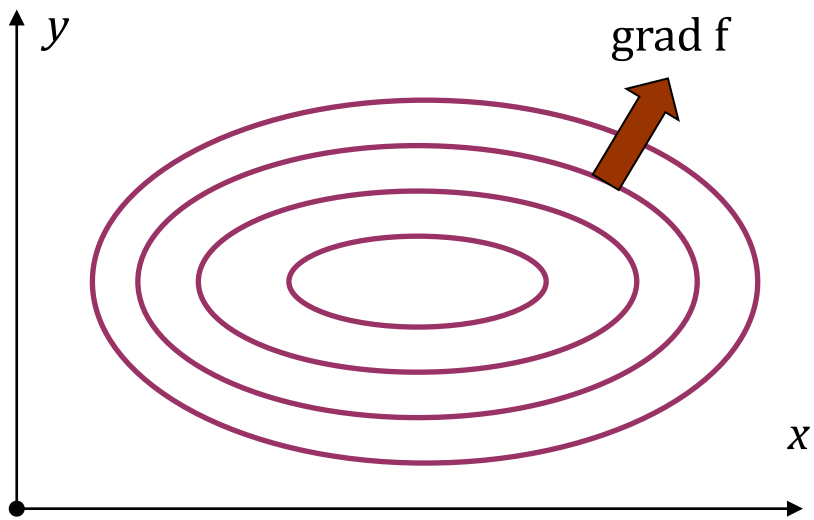


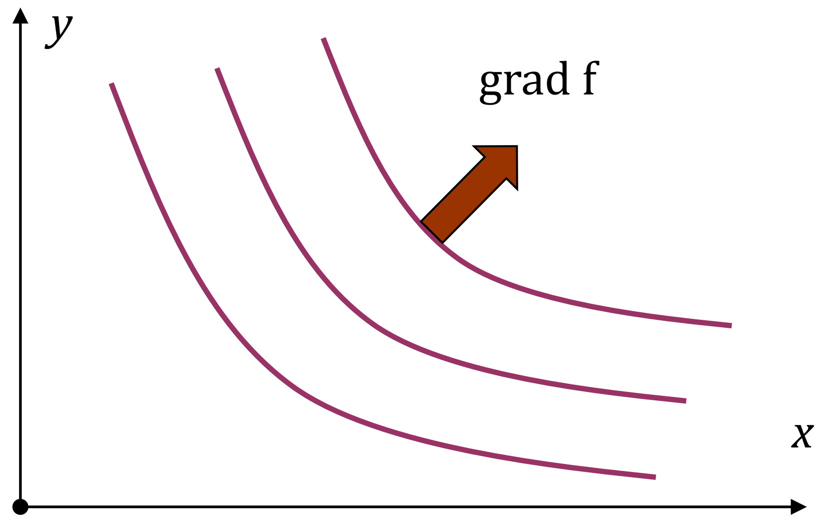
- Постановка задачи:
$$f(x, y) \to \mathrm{extr}$$
- Необходимое условие экстремума:
$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$$
- Матрица Гессе – матрица вторых производных:
$$G = \frac{\partial^2 f}{\partial \mathbf{x}^2} = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}.$$
- Достаточное условие – положительность гессиана:
$$|G| > 0.$$

Целевая функция, линия уровня

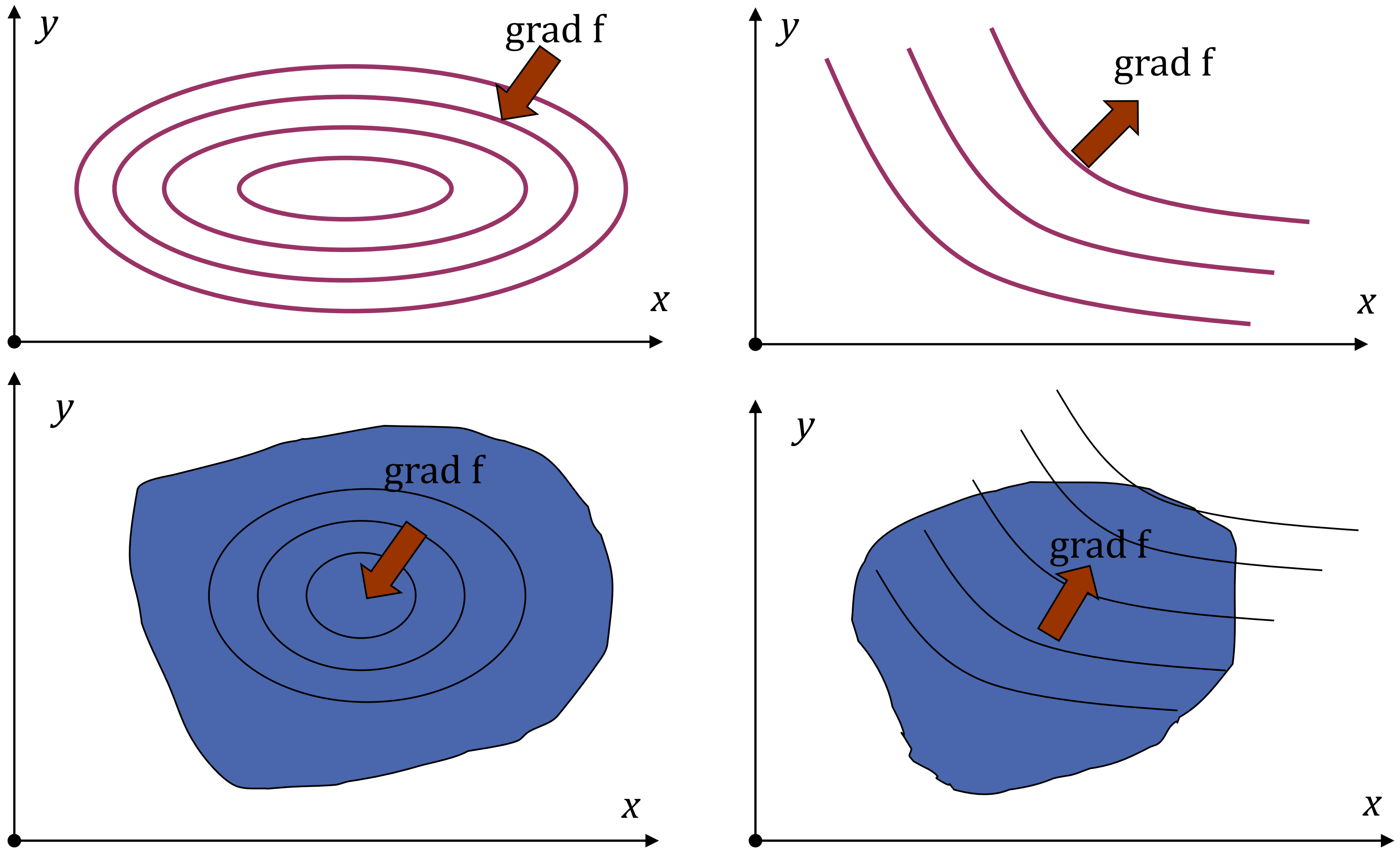
Целевые функции: $f : D \to R, \qquad D \subseteq R^n$

Линии уровня: $f(x, y) = C$.





Линии уровня целевой функции f(x, y) = C

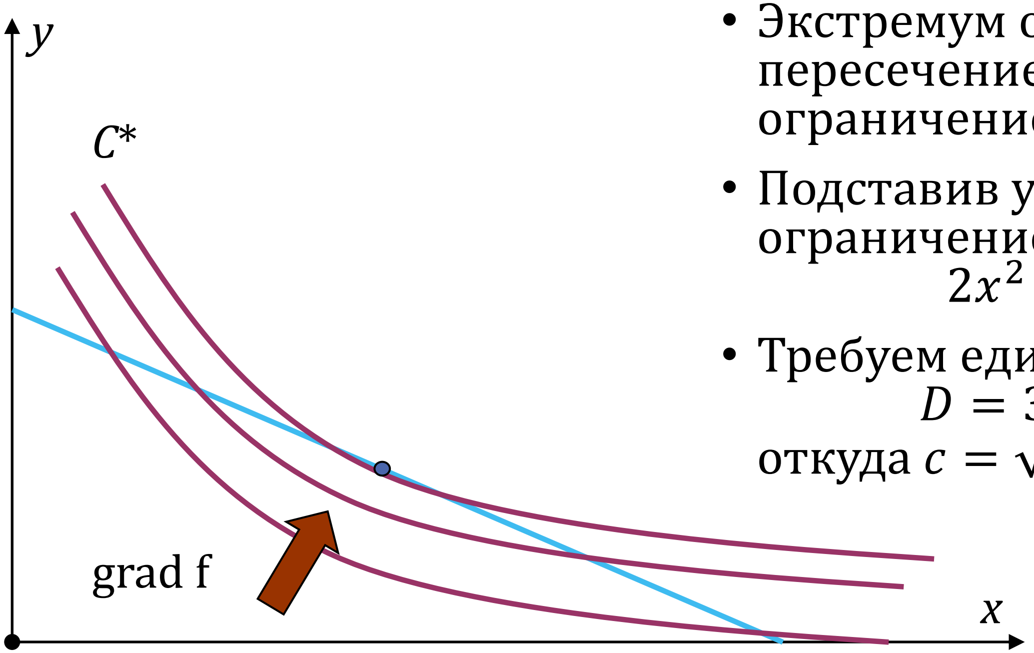


экстремум в стационарной точке

экстремум в граничной точке

Пример задачи с ограничениями

- Пусть требуется решить следующую задачу оптимизации:
$$\sqrt{xy} \to \max.$$
при наличии линейного ограничения:
$$2x + 3y \leq 6.$$
- Линии уровня целевой функции имеют вид $\sqrt{xy} = c$, откуда найдем $x = c^2/y$.



- Экстремум определяется пересечением линии уровня с ограничением.
- Подставив уравнение линии уровня в ограничение, найдем:
$$2x^2 - 6x + 3c^2 = 0.$$
- Требуем единственности корня:
$$D = 36 - 4 \cdot 2 \cdot 3c^2 = 0,$$
откуда $c = \sqrt{6}/2$, x = 3/2, y = 1.

Алгоритмы локальной безусловной оптимизации

- Основная идея ***одношаговых алгоритмов***:
$$\mathbf{X}^{t+1} = (1-\lambda)\mathbf{X}^{t} + \lambda\mathbf{V}^{t+1},$$
где λ – положительная константа, имеющая смысл величины шага в направлении вектора **V**.
- Алгоритмы различаются правилом выбора вектора **V** и числа λ.
- В качестве вектора **V** можно брать градиент, псевдоградиент или случайный вектор.
- Величину шага можно выбирать в несколько этапов. Если при первом выборе не удалось найти решение лучше, можно уменьшить шаг.
- Идея ***многошаговых алгоритмов***:
$$\mathbf{X}^{t+1} = \sum_{\tau=0}^{r-1} \lambda_\tau \mathbf{X}^{t-\tau} + \lambda_V \mathbf{V}^{t+1}, \qquad t > r.$$

Алгоритм имитации отжига

- Пороговый стохастический алгоритм глобальной оптимизации.
- На каждом шаге в окрестности текущего решения выбирается новая точка.
- Принятие этой новой точки происходит с вероятностью:

$$P(X \to X') = \begin{cases} 1, & f(X') \le f(X), \\ \exp\left(-\dfrac{f(X') - f(X)}{T}\right) & f(X') \le f(X). \end{cases}$$

- «Температура» T с каждой последующей итерацией уменьшается по некоторому закону. Например:

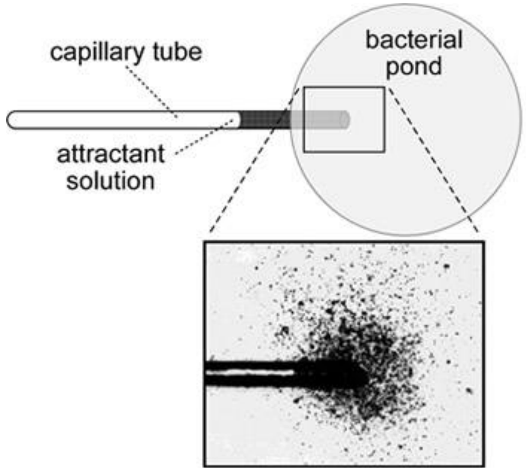
$$T(t) = \alpha\, T(t-1),$$

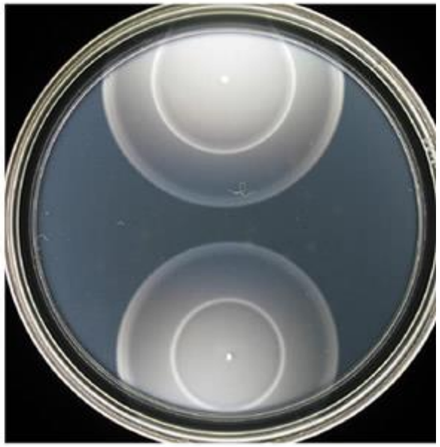
или

$$T(t) = \frac{T_0}{\ln(t+1)}.$$

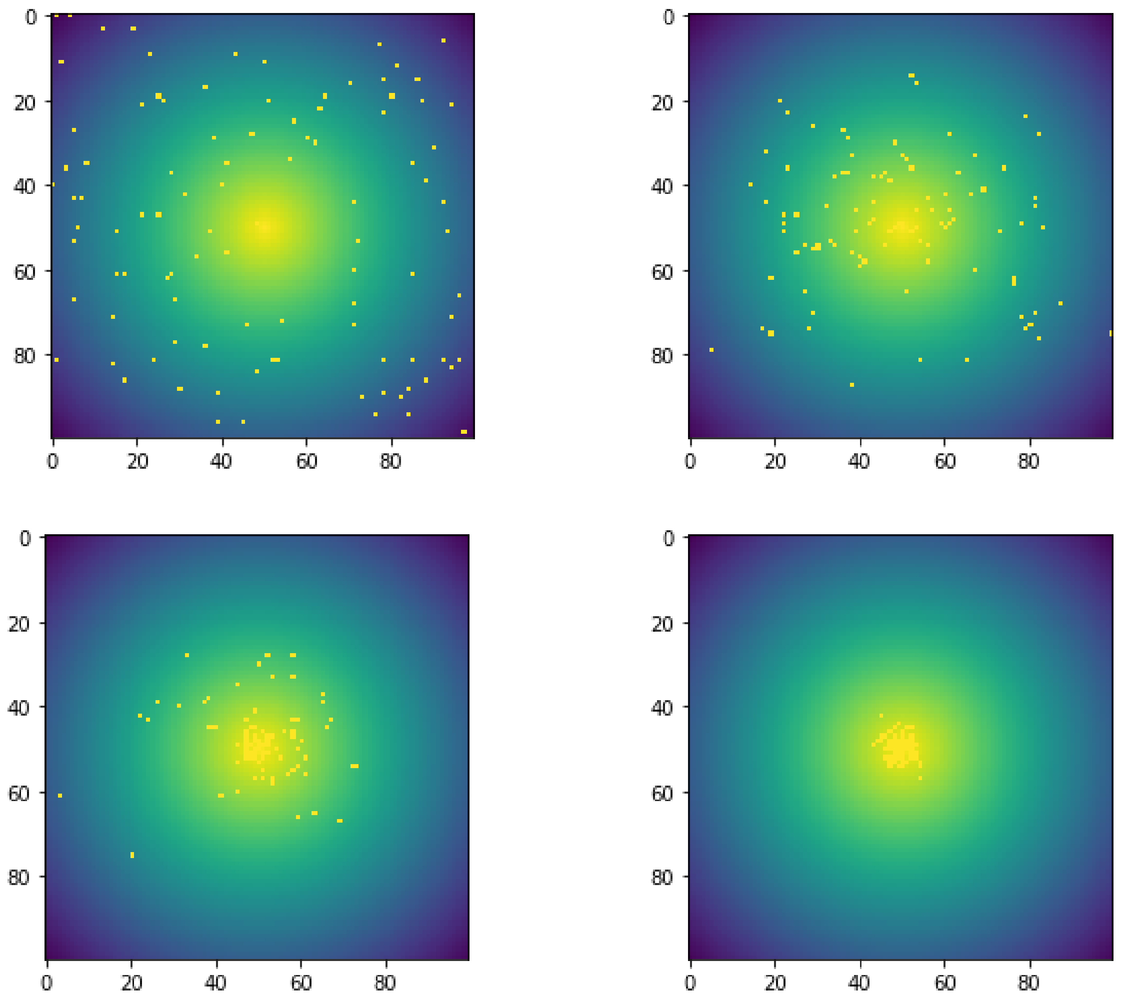
Знакомый пример популяционного алгоритма: хемотаксис бактерий

- Рассматривается эйлерова двумерная сетка.
- В каждой ячейке с координатами (x, y) имеется:
  - список бактерий в ячейке;
  - концентрация питательного вещества $\rho(x,y)$.
- Бактерия – агент i с состоянием (d_i, m_i):
  - d_i – последнее направление перемещения (С, Ю, З, В)
  - m_i – последняя концентрация питательного вещества.
- Бактерия помнит последнюю концентрацию (d_i).
- Бактерия в положении (x, y) чувствует текущую концентрацию $\rho(x, y)$.
- У модели есть два параметра:
  - p_i – вероятность того, что бактерия будет двигаться вперед, если концентрация увеличилась;
  - p_d – вероятность того, что бактерия будет двигаться вперед, если концентрация уменьшилась;
  - Причем, разумеется, $p_i > p_d$.





Результаты оптимизации

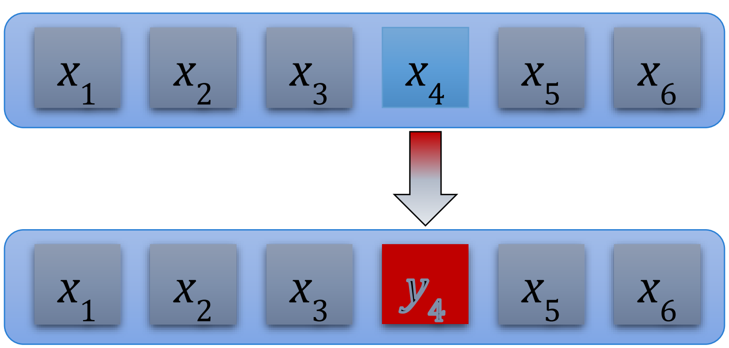


Популяционные алгоритмы: генетический алгоритм глобальной оптимизации

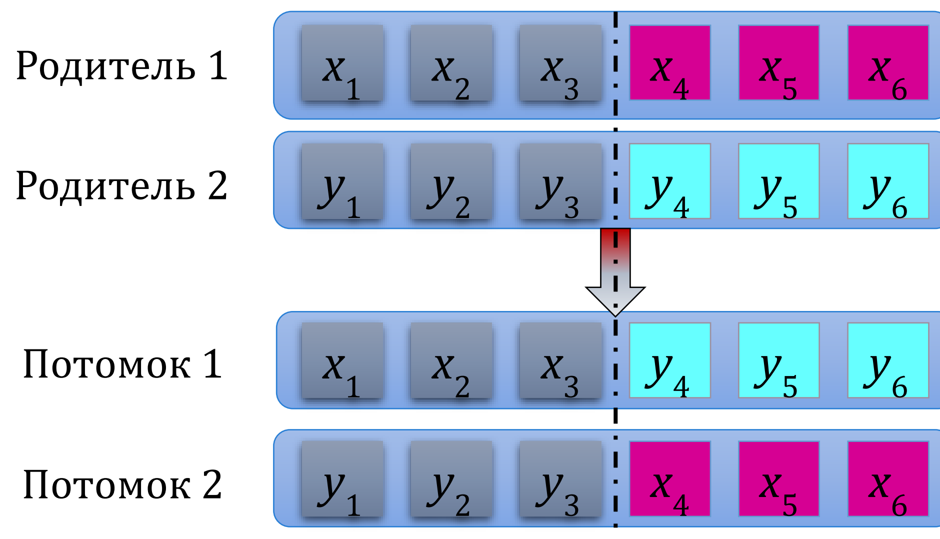
Оператор мутации

Происходит с заданной вероятностью для каждого гена

Случайная точка мутации

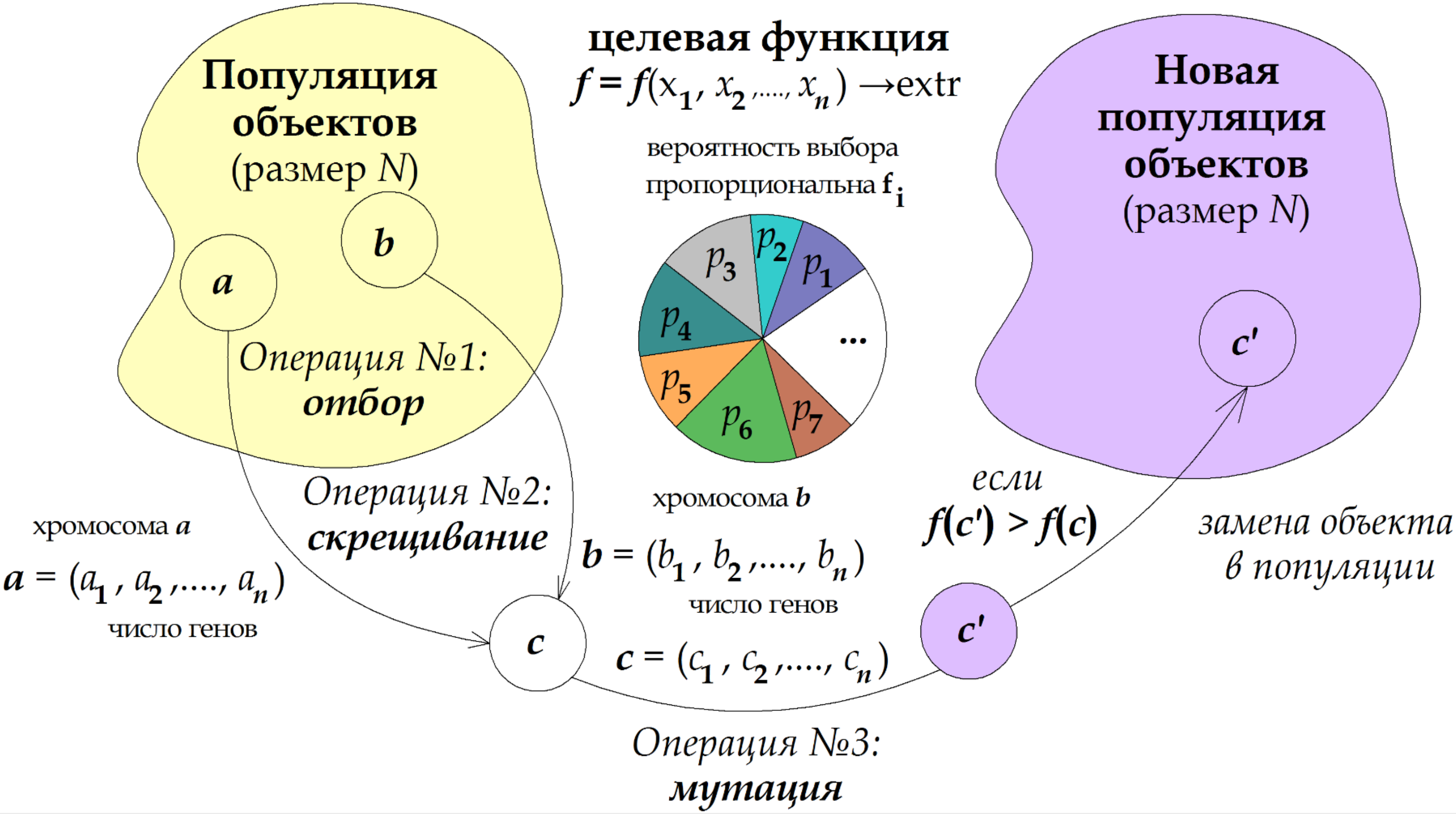


Оператор одноточечного кроссовера



Структура генетического алгоритма

Объекты - хромосомы $\mathbf{x} = (x_1, x_2, \ldots, x_n)$

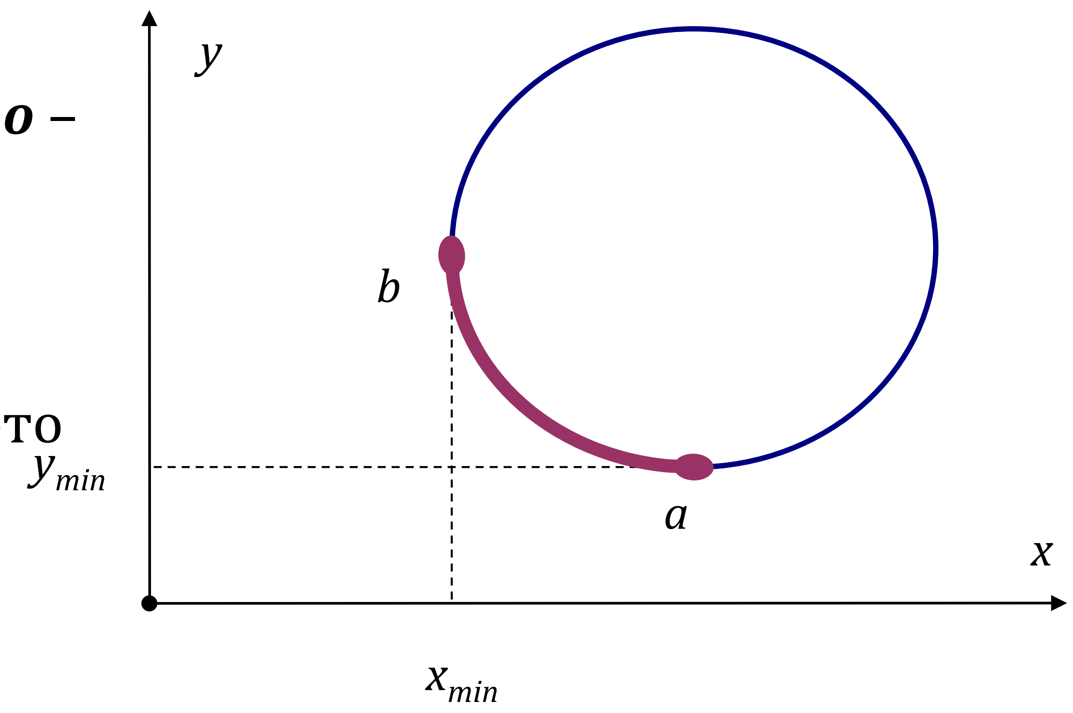


Двухкритериальная задача

- Пусть требуется найти решение следующей задачи оптимизации:

$$\begin{cases} x \to \min \\ y \to \min \end{cases}$$

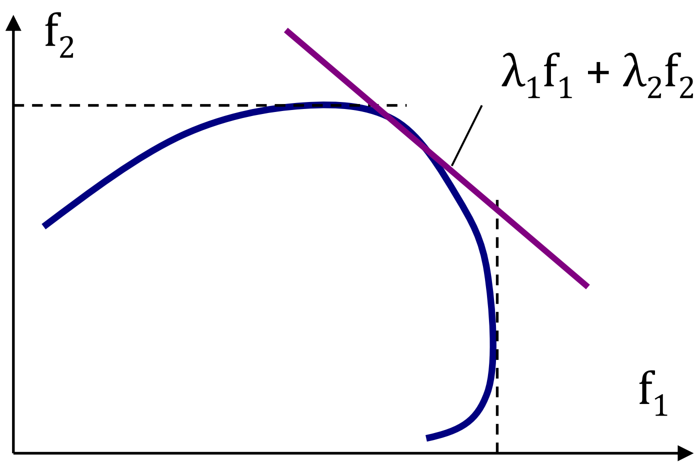
- ***Множество Парето*** – множество недоминируемых альтернатив.
- $\overset{\frown}{ab}$ - множество Парето



Соотношение оценок по критериям

- Оценка альтернативы ω из множества Ω задается векторной оценкой $\mathbf{f} = (f_1, f_2, \dots, f_m)$.
- Доминирование оценки по Парето:
  $\mathbf{f} = (f_1,f_2,\dots,f_m) > \mathbf{g} = (g_1,g_2,\dots,g_m) \Leftrightarrow f_j \geq g_j$.
- Кандидатов на оптимальные решения можно искать только среди парето-оптимальных решений.
- ***Субоптимизация***. Указание нижних границ: для оптимального решения a^* вводится ограничение $f_j(a^*) \geq \gamma_j$ - нижняя граница по j-му критерию. Если f_1 – выделенный критерий, то тогда оптимальной считается такая альтернатива, что на множестве $D' = \{a \in D: f_j(a) \geq \gamma_j, j = 2, \dots, m\}$.
- Если решение не найдено, нужно менять вектор ограничений.
- ***Обобщенный критерий***: $\varphi: Y_1 \times Y_2 \times \dots \times Y_m \to \mathbb{R}$.

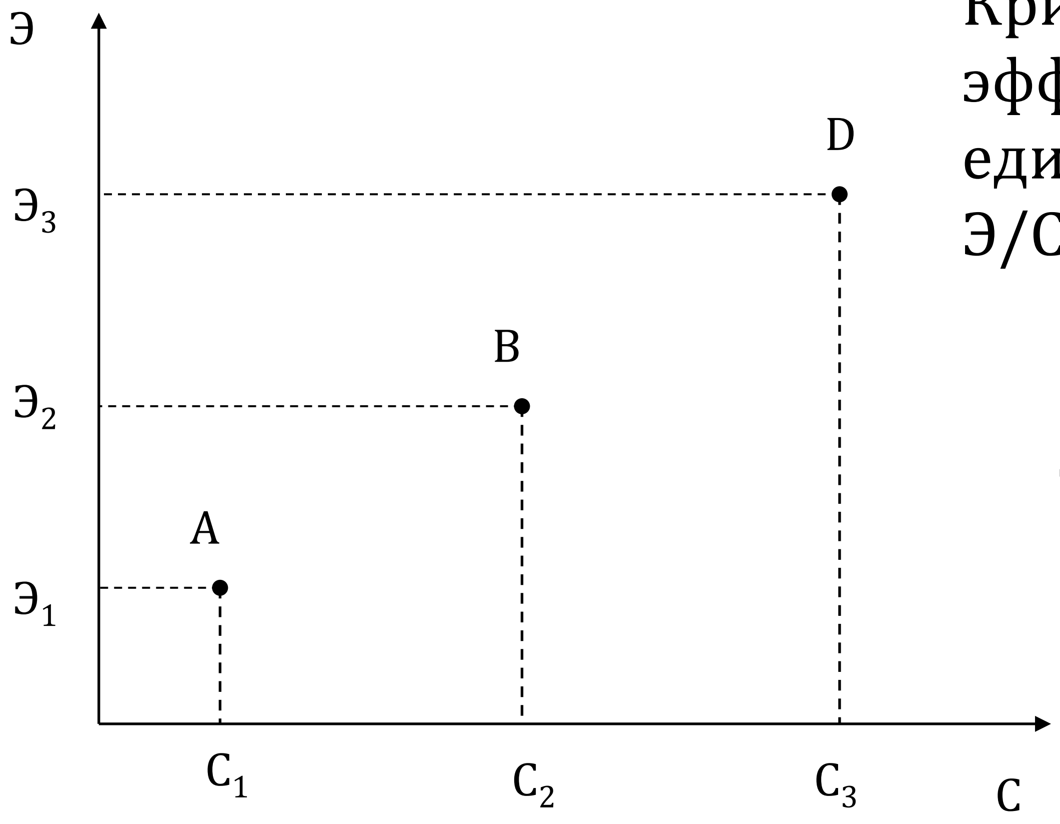
Связь множества Парето и обобщенного критерия



Теорема 1 (прямая). Пусть Q – произвольное множество векторных оценок. Если векторная оценка **f***=$(f_1,\ldots,f_m)$ доставляет максимум функции $\varphi(\mathbf{f^*}) = \alpha_1 f_1 + \ldots + a_m f_m$, где все коэффициенты a_i положительны, то векторная оценка **f*** является Парето-оптимальной в множестве Q.

Теорема 2 (обратная). Пусть Q - выпуклое множество, **f*** - Парето-оптимальная векторная оценка на этом множестве. Тогда найдутся такие неотрицательные числа $a_i \geq 0$, $i=1,\ldots,m$, что функция $\varphi(\mathbf{f^*}) = \alpha_1 f_1 + \ldots + a_m f_m$ достигает максимума на множестве Q в точке **f***.

Критерий стоимость-эффективность



Критерий – удельная эффективность (на единицу стоимости): Э/С

$Э_1/C_1 = Э_2/C_2 = Э_3/C_3$

Независимость предпочтений. Пример

Рассмотрим следующий вариант построения критерия

$$f = \sum_{j=1}^{m} \frac{f_j(a)}{M_j}, \qquad M_j = \max_{a \in D} \left| f_j(a) \right|$$

	з/п (руб)	Отпуск (дни)	Время поездки
A	900	20	-60
B	500	30	-40

$$f(A) = \frac{900}{900} + \frac{20}{30} - \frac{60}{60} = \frac{2}{3};$$

$$f(B) = \frac{500}{900} + \frac{30}{30} - \frac{40}{60} = \frac{8}{9};$$

$$\boxed{B \succ A}$$

	з/п (руб)	Отпуск (дни)	Время поездки
A	900	20	-60
B	500	30	-40
C	400	60	-100

$$f(A) = \frac{900}{900} + \frac{20}{60} - \frac{60}{100} = \frac{22}{30} = \frac{66}{90};$$

$$f(B) = \frac{500}{900} + \frac{30}{60} - \frac{40}{100} = \frac{59}{90};$$

$$f(C) = \frac{400}{900} + \frac{60}{60} - \frac{100}{100} = \frac{4}{9};$$

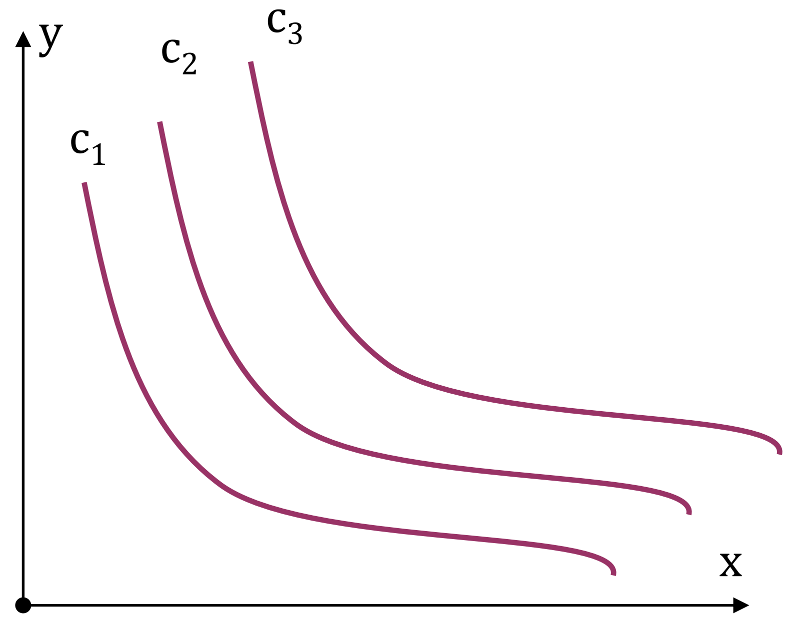
$$\boxed{A \succ B}$$

Обобщенный критерий (другие названия: целевая функция, функция ценности)

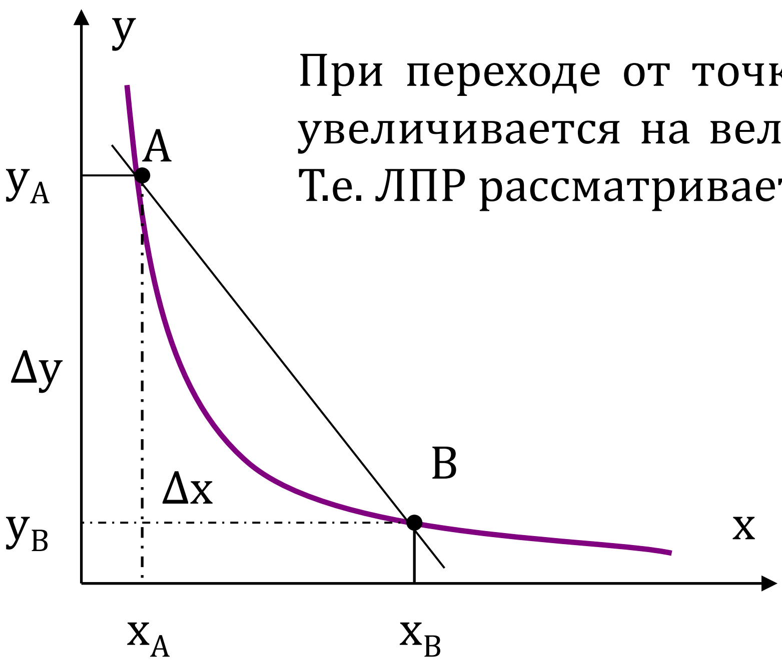
- Основное требование к обобщенному критерию – *сохранение соотношения доминирования по Парето:*
$$a_1 \succ a_2 \quad \Rightarrow \quad \varphi(a_1) > \varphi(a_2).$$
- Это можно рассматривать как определение обобщенного критерия.
- Обобщенные критерии φ_1 и φ_2 эквивалентны (или стратегически эквивалентны), если для <u>любых</u> двух альтернатив a и b выполняется
$$\varphi_1(a) > \varphi_1(b) \quad \Leftrightarrow \quad \varphi_2(a) > \varphi_2(b).$$
- *Пример*: пусть
$$\varphi_1 = 2a + 3b,$$
тогда критерий
$$\varphi_2 = \exp(2a + 3b)$$
эквивалентен φ_1.

Кривая безразличия

- Пусть существует некоторый обобщенный критерий φ(x,y).
- Тогда уравнение φ(x,y) = c определяет при каждом значении константы некоторую кривую на плоскости Oxy.
- Эта кривая называется ***кривой безразличия*** в связи с тем, что вдоль каждой из таких кривых обобщенный критерий принимает одно и то же значение.
- С точки зрения обобщенного критерия все альтернативы, которые лежат на одной и той же кривой равноценны.
- Набор кривых безразличия в криериальном пространстве задает ***карту безразличий***.



Локальный коэффициент замещения



При переходе от точки A к точке B оценка по первому критерию увеличивается на величину Δx, а по второму – уменьшается на Δy. Т.е. ЛПР рассматривает эти оценки как эквивалентные.

Определение. Положительное число

$$k(x,y) = \lim_{\Delta x \to 0}\left(-\frac{\Delta y}{\Delta x}\right) = -\frac{dy}{dx}$$

называется локальным коэффициентом замещения (ЛКЗ) в точке (x,y).

ЛКЗ приблизительно равен той минимальной прибавке по второму критерию, которая компенсирует для ЛПР потерю единицы по первому критерию.

$$k(x,y)dx + dy = 0 \quad \Rightarrow \quad \int k(x,y)dx + y = C \;\Rightarrow\; \varphi(x,y) = \int k(x,y)dx + y$$

В простейшем случае $\quad \Delta y = -k \cdot \Delta x$

Условие постоянства ЛКЗ

Если ЛКЗ постоянен в двумерной области, то дифференциальное уравнение, определяющее его имеет вид

$$\frac{dy}{dx} = -k \quad \Rightarrow \quad y = -kx + C$$

И обратно. Пусть обобщенный критерий представим в виде суммы

$$\varphi(x, y) = \lambda_1 x + \lambda_2 y$$

Тогда кривые безразличия этой функции определяются в виде

$$\lambda_1 x + \lambda_2 y = C \quad \Rightarrow \quad y = C / \lambda_2 - \left(\lambda_1 / \lambda_2\right) x$$

$$\boxed{dy / dx = -\lambda_1 / \lambda_2 = const}$$

! Линейная свертка критериев соответствует постоянству ЛКЗ и линиям уровня в виде параллельных прямых. Это может быть очень существенным ограничением.

Аддитивные функции ценности

Предположим, что ЛКЗ может быть представлен в виде $k(x,y)=X(x)/Y(y)$, тогда

$$X(x)dx+Y(y)dy=0 \qquad \Rightarrow \qquad \varphi(x,y)=\int X(x)dx+\int Y(y)dy$$

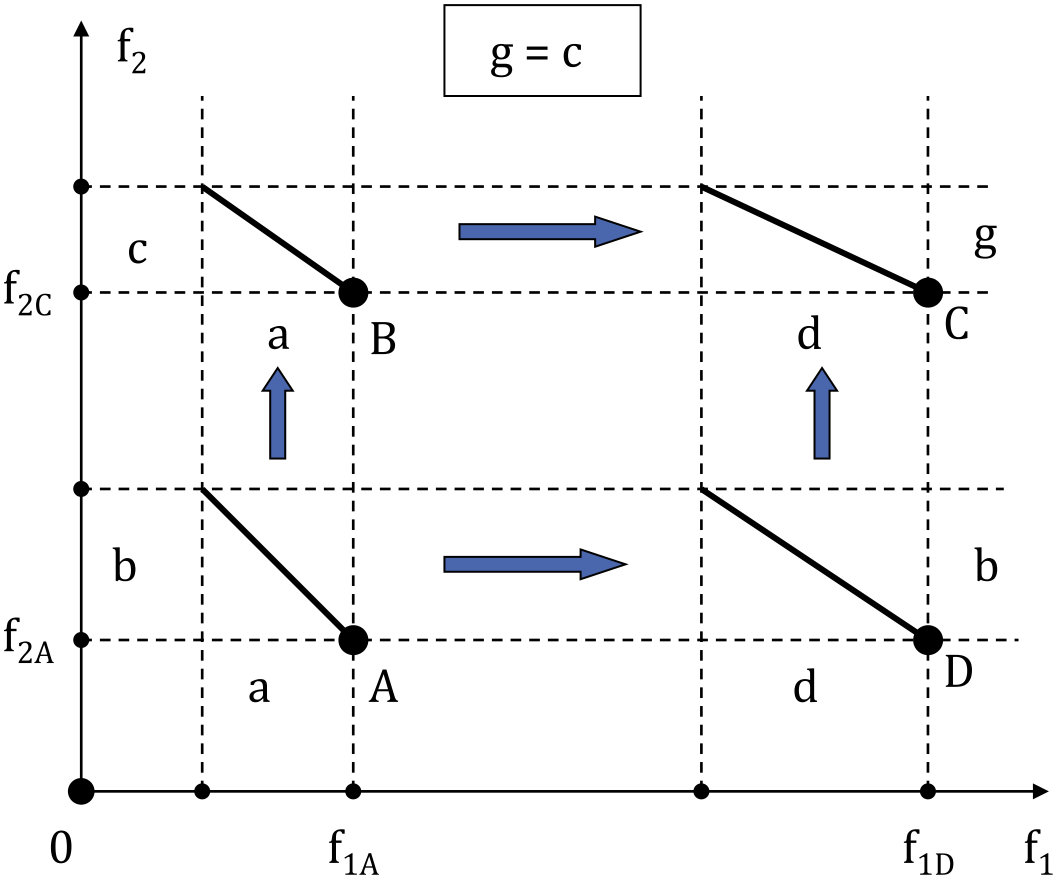
Такой вариант задает аддитивную функцию ценности:

$$\varphi(x,y)=\varphi_1(x)+\varphi_2(y)$$

Или обобщенный вариант $\quad \varphi(f_1, f_2,\ldots, f_m)=\sum_{i=1}^{m}\varphi_i(f_i)$

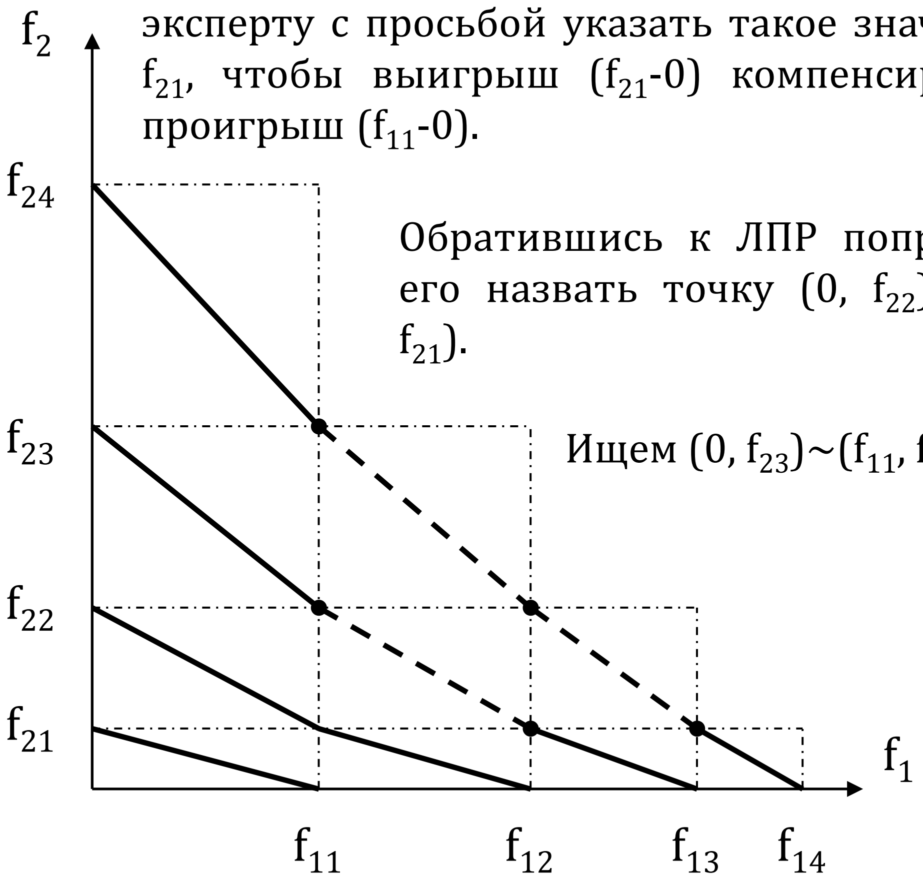
<u>Теорема</u> (Льюис, Тьюки). При наличии двух критериев структура предпочтений аддитивна тогда и только тогда, когда выполняется условие соответственных замещений.

Условие соответственных замещений



Если в т. A при $f_1=f_{1A}$, $f_2=f_{2A}$ потеря a единиц в критерии 1 компенсируется выигрышем b единиц в критерии 2 и если при этом в точке C при $f_1=f_{1C}$, $f_2=f_{2C}$ выигрыш b единиц в критерии 2 может быть замещен потерей d единиц критерия 1, а также в точке B при $f_1=f_{1B}$, $f_2=f_{2B}$ за потерю a единиц в критерии 1 необходимо получить выигрыш в c единиц критерия 2, то для аддитивной структуры предпочтений при $f_1=f_{1D}$, $f_2=f_{2D}$ потеря в d единиц в критерии 1 должна компенсироваться выигрышем в g=c единиц в критерии 2. Если в точках A, B, C известна информация о замещениях критериев 1 и 2, то в точке D она может быть определена для аддитивной структуры предпочтений однозначно.

Построение линий безразличия



Пусть $\varphi(f_{11}, 0) = 1$. Сформулируем вопрос эксперту с просьбой указать такое значение f_{21}, чтобы выигрыш $(f_{21}-0)$ компенсировал проигрыш $(f_{11}-0)$.

Обратившись к ЛПР попросим его назвать точку $(f_{12}, 0) \sim (f_{11}, f_{21})$.

Обратившись к ЛПР попросим его назвать точку $(0, f_{22}) \sim (f_{11}, f_{21})$.

Дальнейшее условие $(f_{13}, 0) \sim (f_{12}, f_{21})$.

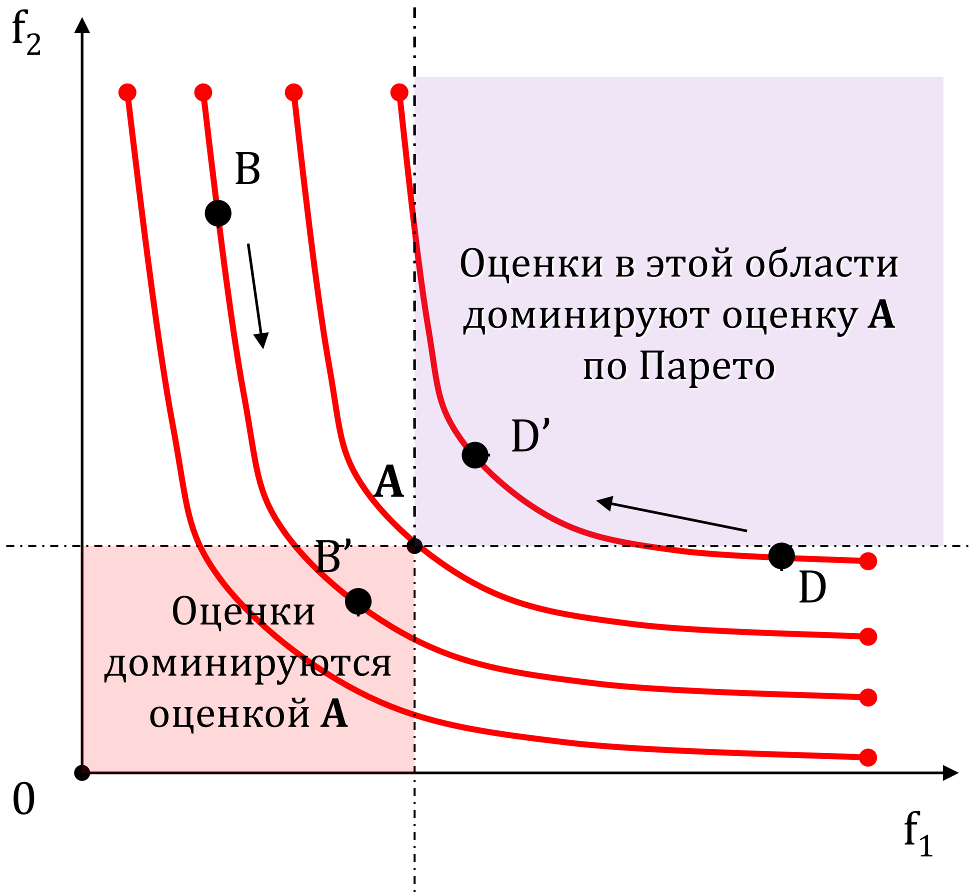
Ищем $(0, f_{23}) \sim (f_{11}, f_{22})$.

Ищем $(f_{14}, 0) \sim (f_{13}, f_{21})$.

И еще $(0, f_{24}) \sim (f_{11}, f_{23})$.

Остальные линии – по условию соответственных замещений.

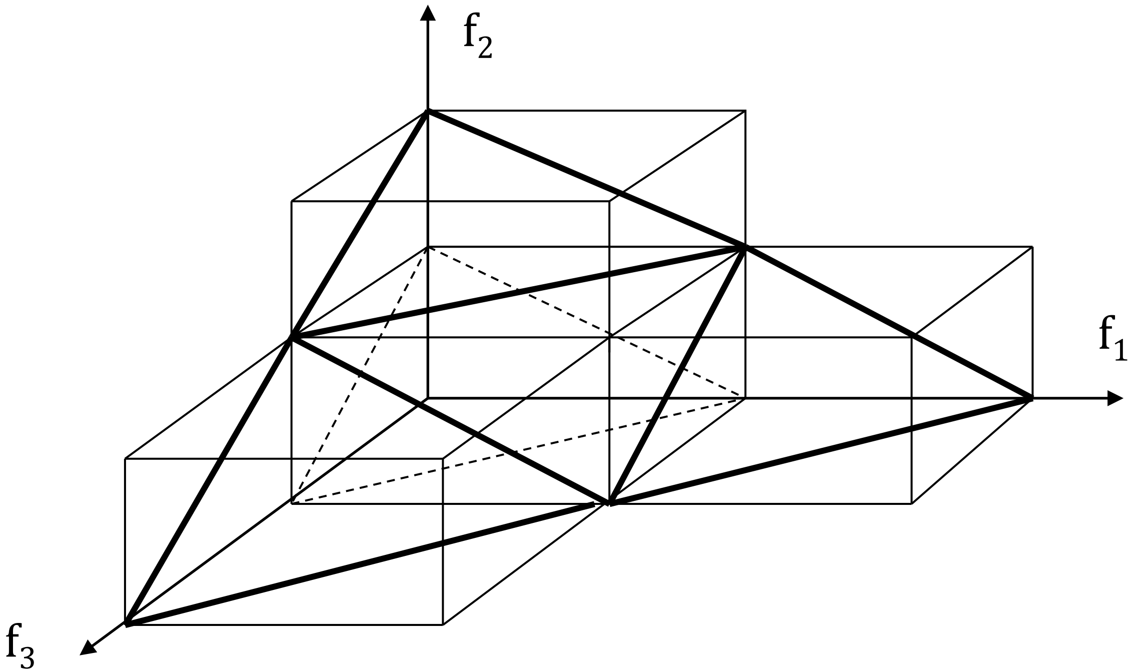
Принятие решений при наличии карты безразличий



- Рассмотрим произвольную точку A на некоторой кривой безразличия.
- Все точки, находящиеся справа и сверху от точки A доминируют ее по Парето.
- Все точки, находящиеся слева и снизу от точки A доминируются по Парето точкой A.
- Такое соотношение можно найти для любых двух точек в силу того, что все точки находящиеся на одной кривой безразличия одинаковы.

Случай трех критериев

- Аддитивная функция ценности:
$$\varphi(f_1, f_2, f_3) = \varphi_1(f_1) + \varphi_2(f_2) + \varphi_3(f_3).$$
- Условие независимости по предпочтениям:
  - Пара <f_1, f_2> не зависит по предпочтению от f_3;
  - Пара <f_1, f_3> не зависит по предпочтению от f_2;
  - Пара <f_2, f_3> не зависит по предпочтению от f_1;



Случай многих критериев

Агрегатирование критериев