



Задачи оптимального управления в динамике генных сетей

Безматерных К.Д.



План

1. Математическая постановка задачи управления
2. Постановки задач ОУ в генных сетях
3. Применение ОУ
4. Методы решения задачи управления
5. Методы оптимизации
6. Задача на модели биосинтеза холестерина
7. Задача на модели внутриорганизменной инфекции туберкулеза
8. Задача на модели эпидемии туберкулеза



Постановка задачи оптимального управления (ОУ)

$$\dot{x} = F(t, X, U, P) (*)$$

t – независимая переменная (время).

$X = (x_1 \dots x_N)$ – N -мерный вектор фазовых координат.

$U = (u_1 \dots u_M)$ – M -мерный вектор управляющих функций.

$F = (f_1 \dots f_N)$ – N -мерный вектор функция.

$P = (p_1 \dots p_K)$ – K -мерный вектор параметров.

На фазовые координаты и управляющие функции наложены ограничения:

$X(t)$ принадлежит $G(t)$,

$U(t)$ принадлежит $G_u(t, x)$ ($0 \leq t \leq T$) (**)

Найти такую фазовую траекторию $x(t)$ и соответствующее ей управление $U(t)$ на интервале $[0, T]$, чтобы удовлетворить системе (*), ограничениям (**) и минимизировать функционал J .



Постановка задачи ОУ в генных сетях

- **Задача восстановления функционирования генной сети (ГС)** - задача восстановления *нормального* стационарного состояния или *нормальной* динамики поврежденной (мутантной) ГС при заданных ограничениях на переходной процесс, при минимизации ограничений и/или минимизации некоторого функционала.
- **Задача оптимизации ГС** - Задача поиска необходимых воздействия на ГС для оптимизации наработки некоторого вещества.
- **Ограничения** - ограничения на концентрации веществ, на время переходного процесса, температуру и др.
- **Функционал** - затраты лекарственных препаратов, энергии, времени и др.



Актуальность задач ОУ в генных сетях

- Задача восстановления функционирования ГС – разработка современных лекарственных препаратов
- Задачи оптимизации ГС – создание высокоэффективных продуцентов



Методы решения задач ОУ

- Принцип максимума Понтрягина -> краевая задача
- Принцип максимума Беллмана -> перебор по всем возможным состояниям фазовых переменных



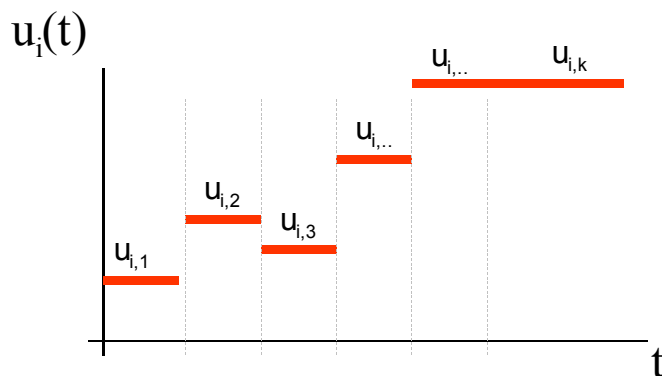
Численные методы решения задач ОУ

- **Метод Сеток** (сведение к задаче поиска минимума конечномерной функции с ограничениями)
- **Метод Пробных Функций** (сведение к задаче поиска минимума конечномерной функции с ограничениями)
- *Метод Локальных Вариаций*
- *Динамическое программирование*

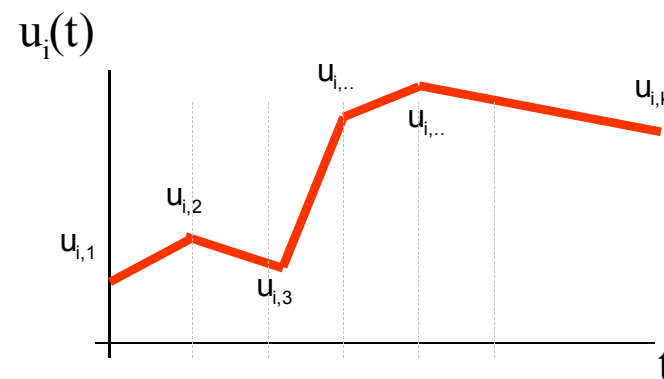


Метод сеток

Суть метода: Сведение к задаче минимизации конечномерной функции с помощью параметризации управляющих функций на конечном интервале времени



кусочно-постоянное управление



кусочно-линейное управление

$$u_i(t) \leftrightarrow u_i(u_{i,1}, u_{i,2}, \dots, u_{i,k})$$

$$U(t) = (u_1(t), u_2(t), \dots, u_m(t)) \leftrightarrow (u_1(u_{1,1}, u_{1,2}, \dots, u_{1,k}), \dots, u_m(u_{m,1}, u_{m,2}, \dots, u_{m,k}))$$



Метод пробных функций

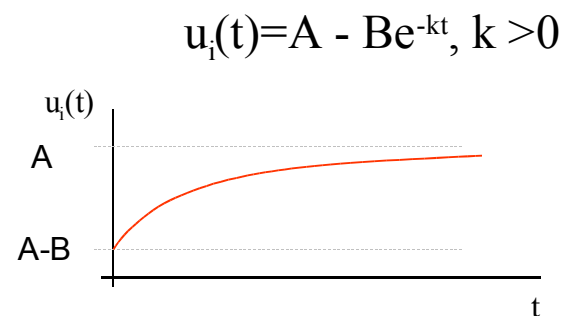
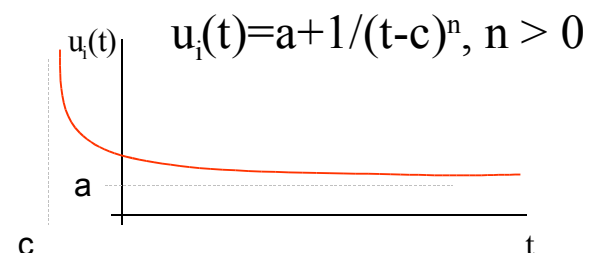
Суть метода: Сведение к задаче минимизации конечномерной функции используя пробные функции в качестве управляющих

1. $\{v_n\}$ – полная система функций
2. $J(v_n)$ – непрерывный функционал

$v_n(x; a_1^* \dots a_n^*)$ – последовательность минимизирующих функций, $a_1^* \dots a_n^*$ – решение задачи минимизации функционала J на n -ом шаге.

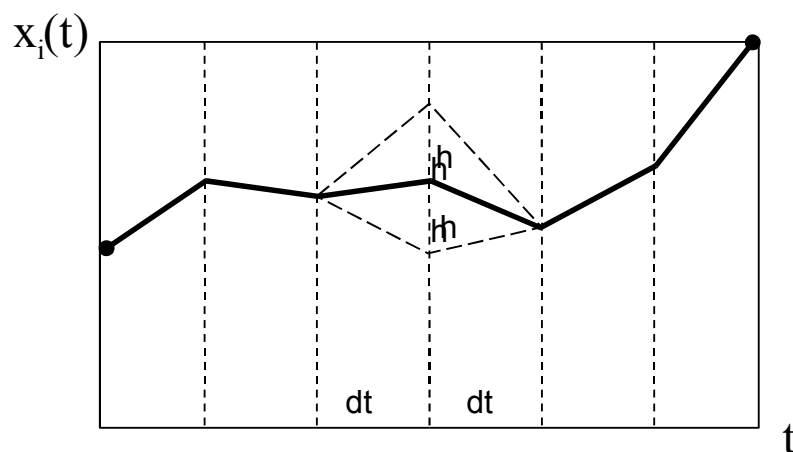
$$v_n(x; a_1 \dots a_n) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n$$

$$v_n(x; a_1 \dots a_n) = a_0 + a_1 \sin(x) + a_2 \sin(2x) + \dots + a_n \sin(nx)$$





Суть метода: Варьирование фазовых переменных на сетке, на каждом шаге алгоритма решение элементарной задачи управления.

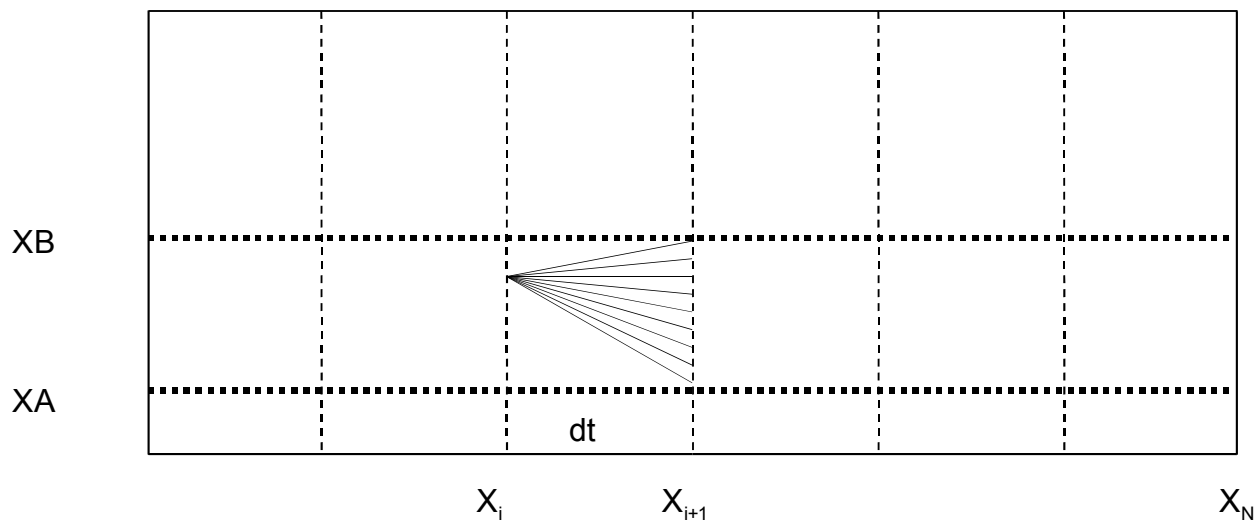


Элементарная операция – поиск допустимого оптимального управления из точки (t_1, x_1) в точку (t_2, x_2) , причем $t_2 - t_1 = dt$.



Метод динамического программирования

Суть метода: Метод динамического программирования состоит в переборе всех возможных состояний фазового вектора X в «трубке» ($X_A; X_B$) при фиксированных dx и dt .



Метод обладает глобальной сходимостью в «трубке», но требует:

памяти $\sim N \cdot S \cdot n \cdot 8$ байт, где N – количество интервалов по времени, S – ширина трубки (сетка по x), n – число переменных управления, 8 – размер типа `double` в C/C++.

Времени $\sim N \cdot (S \cdot n)^2 \cdot \delta$, где δ – время на приближенное решение краевой задачи (при заданных X_i , X_{i+1} , t_i и t_{i+1} восстановить управление u).



Методы оптимизации

- Одномерная оптимизация
 - Метод деления пополам
 - Метод золотого сечения
 - Метод парабол, метод Брента
- Многомерная оптимизация
 - Методы без производных
 - покоординатный спуск
 - Метод Хука-Дживса
 - Метод Розенброка ...
 - Градиентные методы
 - Метод наискорейшего спуска
 - методы сопряженных направлений
 - методы Ньютона
 - квази-ньютоновские методы (BFGS)
- Поиск глобального минимума
 - Детерминированные
 - Стохастические
 - Метод мултистарта
 - Метод ветвей и границ (интервальные методы)
 - Метод обжига
 - ЭВ, ГА
- Учет ограничений
 - Метод штрафных функций



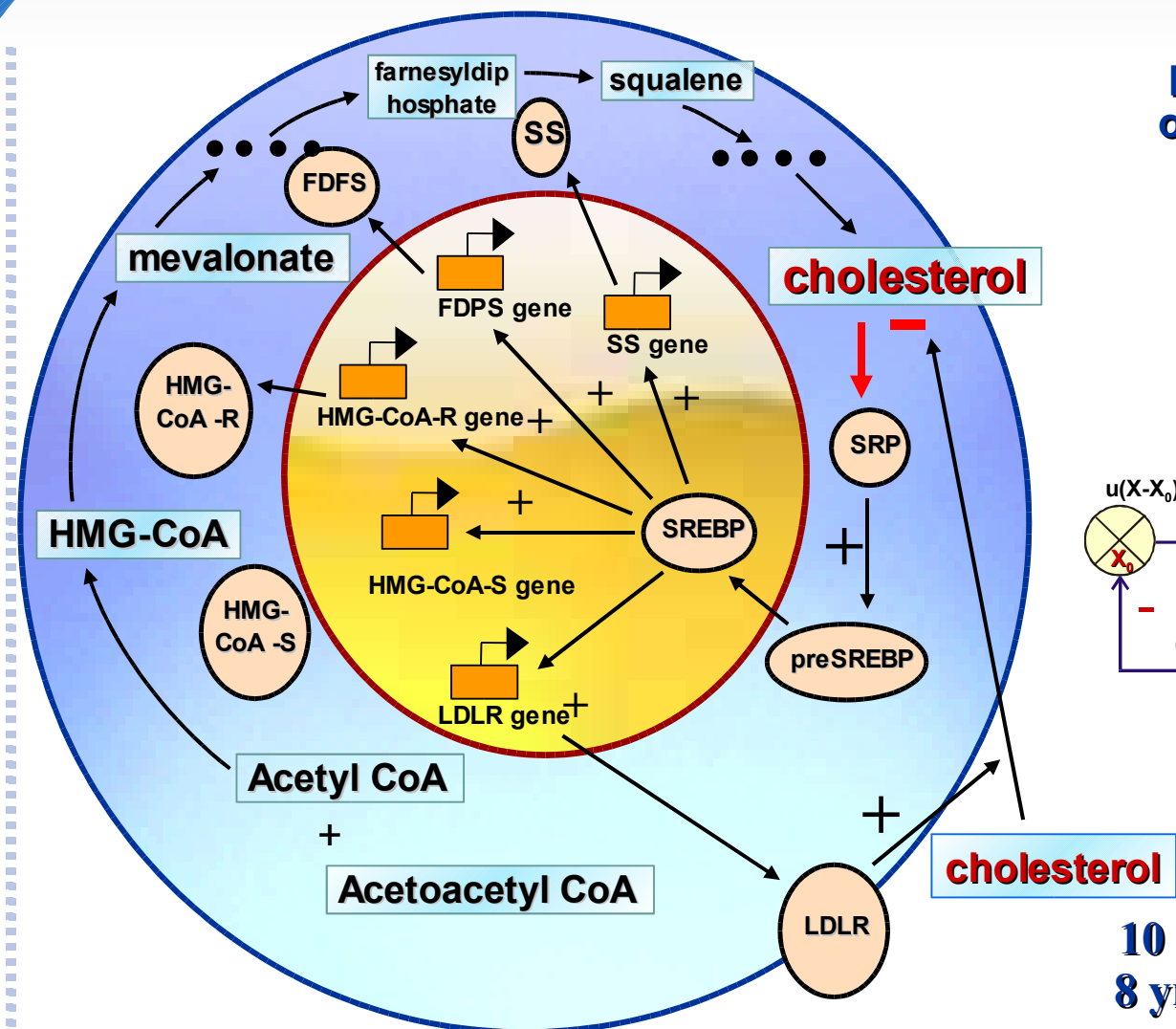
Пакет «ПОИСК»

1. Два покоординатных спуска со случайным выбором координат
2. Определение вектора уменьшения функции X
3. Сканирующий конус вдоль вектора X
4. Случайные шары
5. Уточнение минимума
6. Повторение шагов 1-5

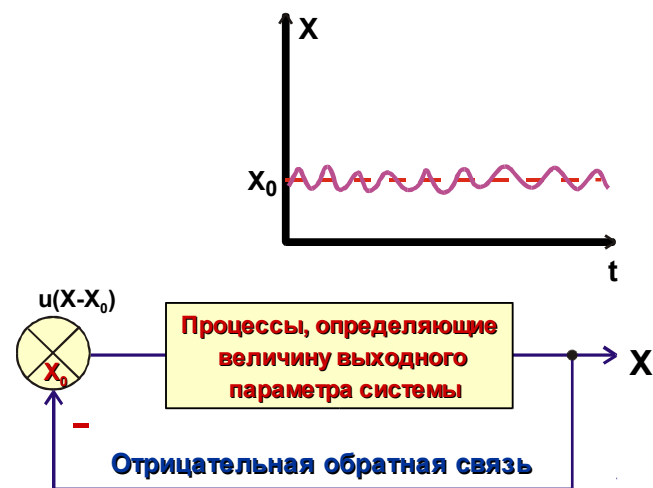
8. Адаптация параметров в процессе поиска (локализация минимума, угол раствора конуса)



База данных **GENENET**: центральный фрагмент геной сети биосинтеза холестерина в клетке (регуляция по механизму отрицательной обратной связи)



Принципиальная схема регуляторного контура с отрицательной обратной связью



10 фазовых переменных
8 управляемых констант



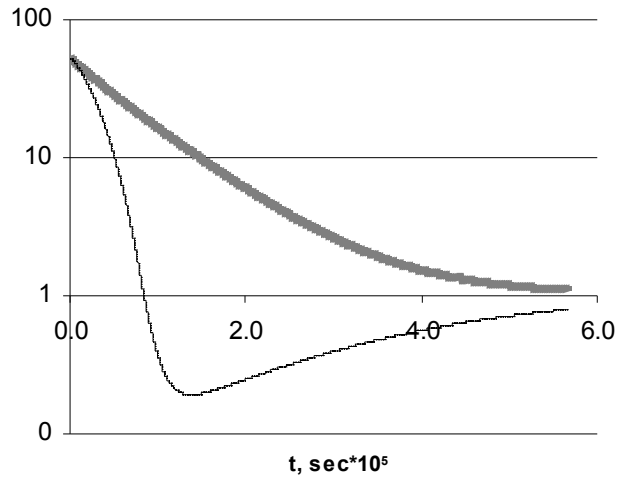
Задача на упрощенной модели биосинтеза холестерина(1)

- Мутация задана изменением 2-х констант
- 2. Восстановить стационарное состояние системы в окрестности нормы (минимум энергетических затрат).
- 3. Минимизировать время переходного процесса + удержание в окрестности нормы.

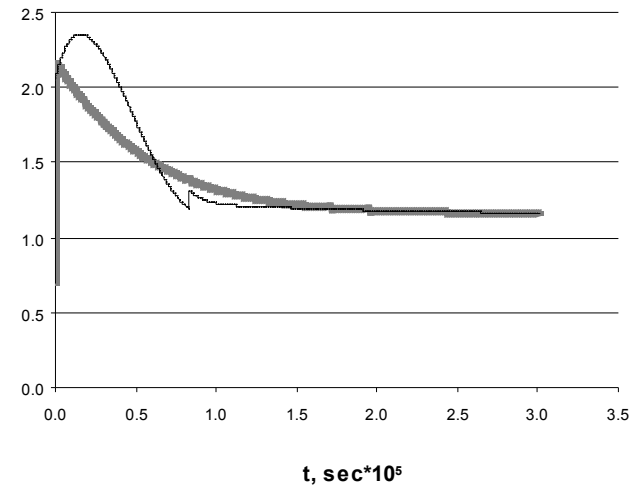


Задача на упрощенной модели биосинтеза холестерина(2)

low-density lipoproteins



cholesterol ether





Модель инфекции туберкулеза(1)

Переменные модели	Пояснения
Базовая модель	
$B_E(t)$	Концентрация внеклеточных бактерий
$B_I(t)$	Концентрация внутриклеточных бактерий
$M_R(t)$	Концентрация покоящихся макрофагов
$M_A(t)$	Концентрация активированных макрофагов
$M_I(t)$	Концентрация инфицированных макрофагов
$T_{pl}(t)$	Концентрация предшественников CD4+ Т-лимфоцитов в легких
$T_1(t)$	Концентрация Т-лимфоцитов-хелперов 1-го типа (Th1)
$T_2(t)$	Концентрация Т-лимфоцитов-хелперов 2-го типа (Th2)
$I_\gamma(t)$	Концентрация IFN- γ
$I_4(t)$	Концентрация IL-4
$I_{10}(t)$	Концентрация IL-10
$I_{12}(t)$	Концентрация IL-12 в легких
$Iln_{12}(t)$	Концентрация IL-12 в лимфоузле
$T(t)$	Концентрация наивных Т-лимфоцитов
$T_{ph}(t)$	Концентрация предшественников CD4+ Т-лимфоцитов в лимфоузле
$MDC(t)$	Концентрация зрелых дендритных клеток в лимфоузле
$IDC(t)$	Концентрация незрелых дендритных клеток в легких

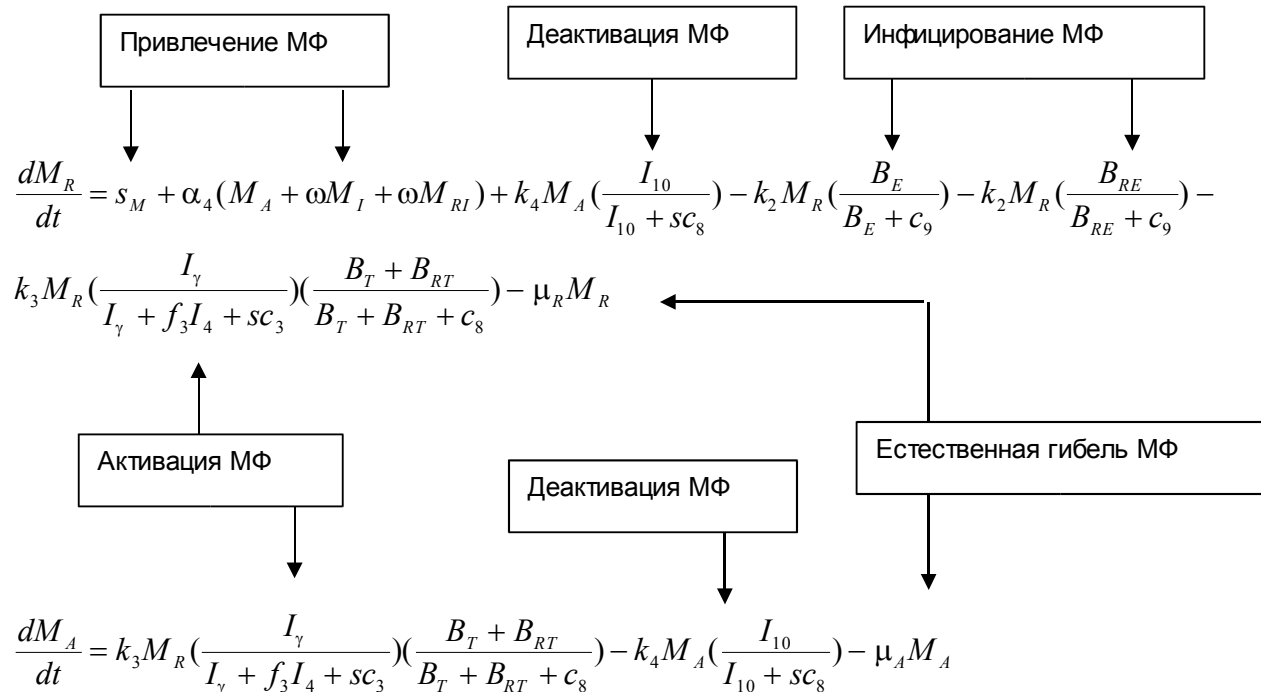
Отличие модели от существующих:

2. Модель учитывает возможность лечения
3. Учет возникновения лекарственно-устойчивых штаммов
4. Учет воздействия лекарств на печень больного.



Модель инфекции туберкулеза(2)

Динамика макрофагов:



...



Модель инфекции туберкулеза(3)

Задача ОУ

Характеристики интервенции (схемы лечения)	Параметры модели	Диапазон
Перечень выбранных препаратов	$li, i=1,2,3,\dots,9$	$li \in [0/1]$
Дозы выбранных препаратов	$sDi, i=1,2,3,\dots,9$	$0 < sDi < \max(sDi)$
Время (длительность) применения выбранных препаратов	$ti, i=1,2,3,\dots,9$	$0 < ti < \max(ti)$
Применение гепатопротекторов	lp, sP	$lp \in [0/1]$

Задача управления заключается в оптимизации схемы лечения при заданных начальных данных и параметрах модели, а также при некоторых ограничениях.

Оптимизирующие критерии:

Эффективность лечения E

Степень поражения печени ΔH

Обобщенный критерий оптимизации K

Учет ограничений:

$0 \leq \Delta H < P,$

$T < T_{max},$

Появление лекарственно устойчивых форм



Модель эпидемии туберкулеза(1)

$$dX/dt = \pi X - \sum \beta_i(X) * (T_i + m_i * T_{0i}) - \mu X$$

$$dR_i/dt = \pi R_i + (r_i + s_i) * L_j + (r_{i0} + s_{i0}) * L_{0j} + (1-p) * \beta_i(X) * (T_i + m_i * T_{0i}) - (d_i * \text{inf} + \mu + \delta_i) * R_i$$

$$dL_i/dt = \pi L_i + (p_i * \beta_i(X) + r_i + f_i + d_i * \sum \beta_i(R_j)) * (T_i + m_i * T_{0i}) - (k_i + v_i + \mu + \mu l_i + r_l i) * L_i + \delta_i * R_i$$

$$dL_{0i}/dt = \pi L_{0i} + (r_{t0i} + f_{0i}) * T_i + k_i * L_i - (v_{0i} + \mu + \mu l_i + r_{l0i} + s_i) * L_i + \sum \gamma_{ji} * L_j$$

$$dT_i/dt = \pi T_i + v_i L_i - (k_t i + \mu + \mu t_i + r_t i + f_i) T_i$$

$$dT_{0i}/dt = \pi T_{0i} + v_{0i} * L_{0i} + k_t i * T_i - (\mu + \mu t_i + r_{t0i} + f_i) T_{0i} + \sum \eta_{ji} * T_{0j} dt$$

$$\beta_i(X) = \beta_{0i} * X / (N + A), \quad \beta_i(R_j) = \beta_{0i} * R_j / (N + A)$$

$$\text{inf} = \sum \beta_{0i} / (N + A) * (T_i + m_i * T_{0i})$$

$$\eta_{0i} = \Theta_{0i} * G_{t0} / Q_{t0}, \quad \gamma_{0i} = \Gamma_{0i} * G_{l0} / Q_{l0} \text{ для } i > 0 \text{ и } \eta_{j0} = \Theta_{j0} * (1 - G_{tj}), \quad \gamma_{j0} = \Gamma_{j0} * (1 - G_{lj})$$

$$s_i = G_{l_i} * Q_{l_i}, \quad f_i = G_{t_i} * Q_{t_i}$$



Модель эпидемии туберкулеза(2)

Задача ОУ

Сумма средств, выделяемая на все интервенции в год, млн.руб.: 50.0		Продолжительность воздействия, лет: 5	Оценка результата через 7 лет	
Интервенция	«Цена»*).	Ожидаемый эффект	Вкл.?	
Увеличение числа койко-мест	0.1	Повышается доля госпитализированных лиц среди больных, нуждающихся в стационарном лечении, и, как следствие, увеличиваются охват лечением и эффективность лечения	✓	
Увеличение адекватности лечения	0.7	Применение таких препаратов и схем лечения, которые увеличивают эффективность терапии, особенно для штаммов с множественной лекарственной устойчивостью	✓	
Ревакцинация здоровых лиц, утративших Поствакцинал ь-ный иммунитет	1.5	Ожидаемый эффект от ревакцинации – снижение случаев инфицирования	✓	
Генотипирование ДНК возбудителя и больного	2.3	Может увеличить скорость дифференциальной диагностики туберкулеза, повысить эффективность лечения за счет быстрого определения индивидуальной скорости метаболизма противотуберкулезных препаратов и выявления наследственно-обусловленных иммунодефицитов.	✓	



Модель эпидемии туберкулеза(3)

Задача ОУ

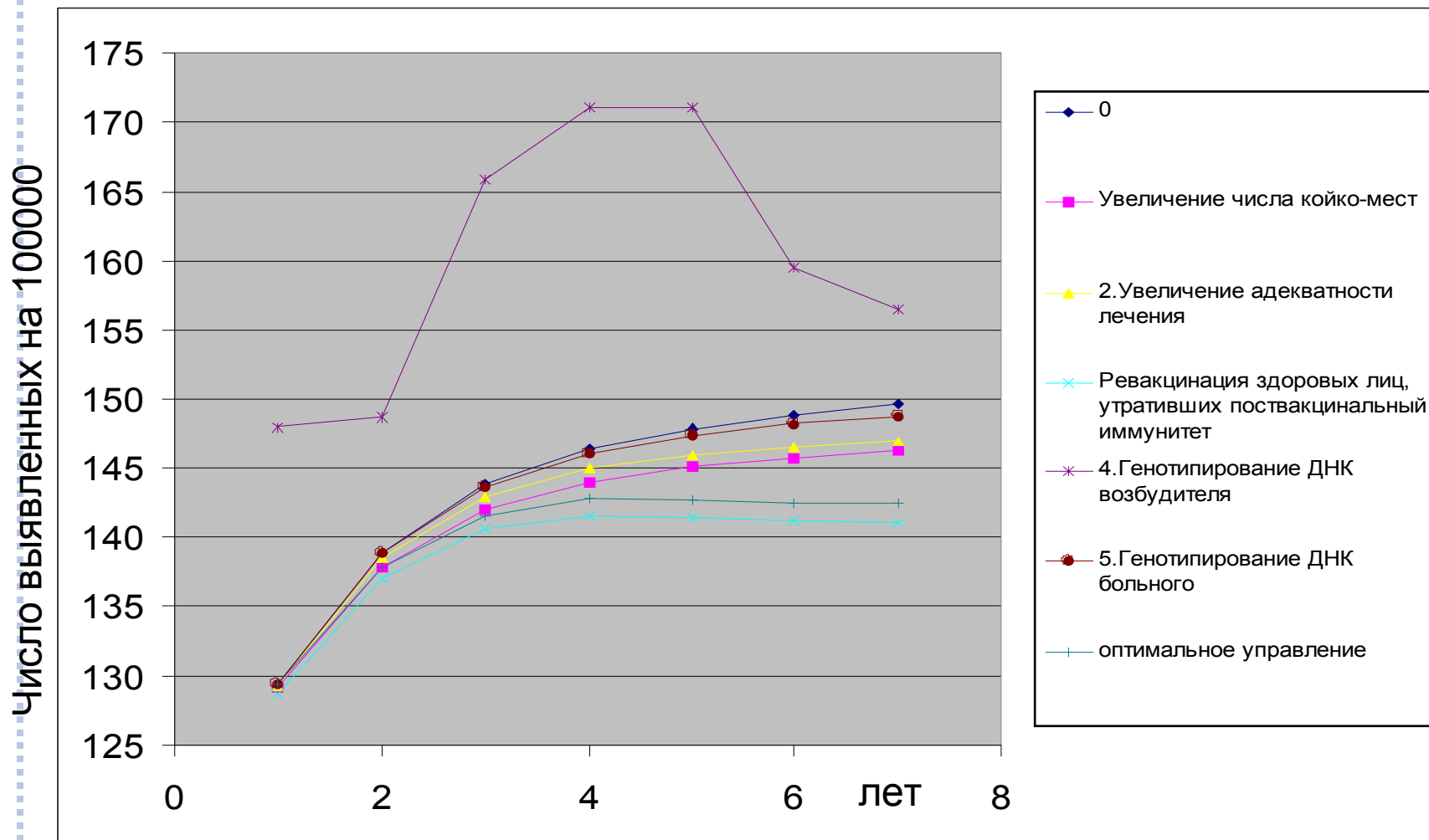
Критерий	Единица эффективности	Выбрать?	Примечания
Заболеваемость через 7 лет (на 100000 населения)	Снижение в процентах к исходному значению		
Заболеваемость в течение 7 лет (на 100000 населения)	Снижение в процентах к исходному значению		
Распространенность через 7 лет (на 100000 населения)	Снижение в процентах к исходному значению	√	
Распространенность в течение 7 лет (на 100000 населения)	Снижение в процентах к исходному значению		
Смертность через 7 лет (на 100000 населения)	Снижение в процентах к исходному значению		
Смертность в течение 7 лет (на 100000 населения)	Снижение в процентах к исходному значению		



Модель эпидемии туберкулеза(3)

Результаты

Оптимизация числа выявленных заболевших через 7 лет

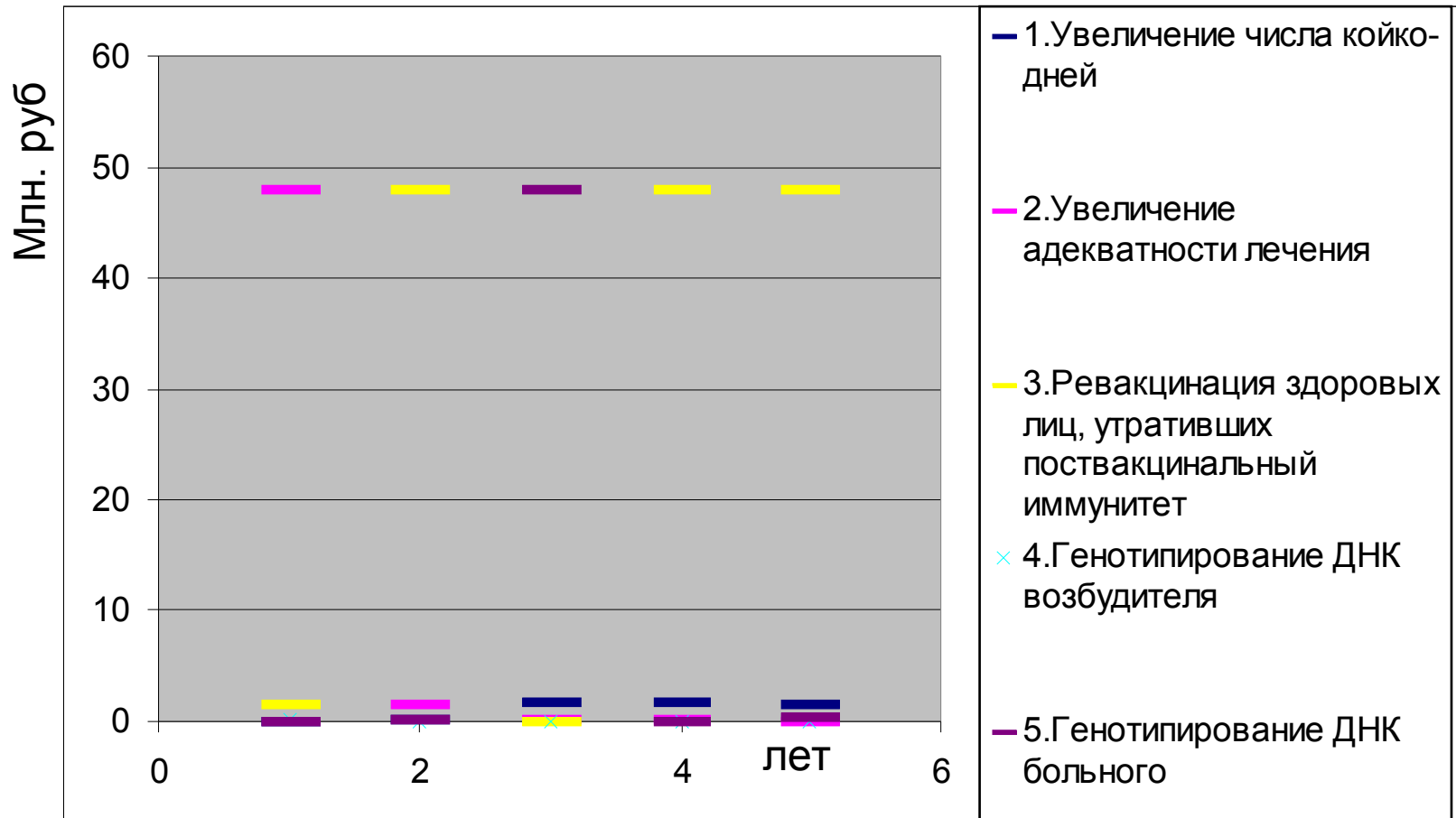




Модель эпидемии туберкулеза(4)

Результаты

Оптимизация числа выявленных заболевших через 7 лет





Задачи ОУ в генных сетях и не только

- восстановление нормальной работы ГС, стационарного состояния или динамики
- создание высокоэффективных продуцентов (наработка необходимых белков)
- управление ходом эксперимента (подбор динамических характеристик)
- управление эпидемиями (экономический аспект)



*Спасибо за
внимание*



Полезные ссылки:

<http://wwwmgs.bionet.nsc.ru/systems/mgl/genenet/> - база данных GenNet

<http://www.math.utk.edu/~lenhart/smb2003.v2.html> - курс лекций и практических занятий по ОУ в биологии (University of Tennessee, Knoxville)

<http://www.mat.univie.ac.at/~neum/glopt.html> - сайт посвященный методам глобальной оптимизации

<http://www.ccas.ru/depart/malashen/53kmsu.htm> - Новикова Н М курс лекций по оптимизации

<http://www.bionet.nsc.ru/chair/cib/cib.php?f=lectures&p=lectures>
Лекции кафедры биоинформатики ФЕН НГУ