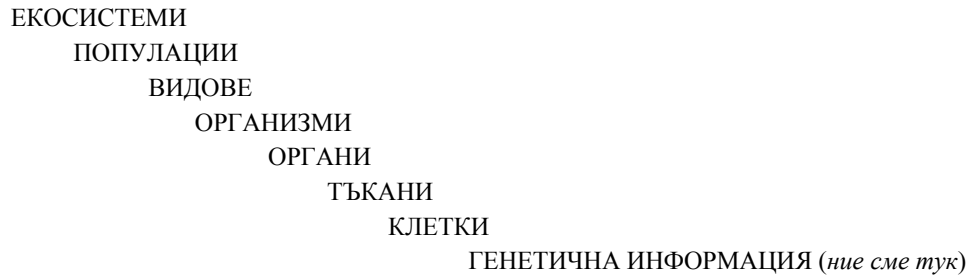
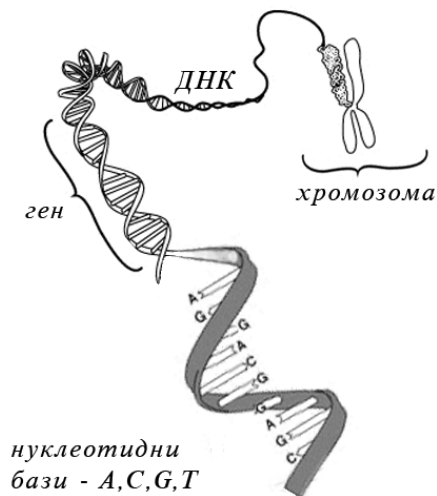


## Общ поглед. Основни модели и понятия

Молекулярната биология, като представител на системната биология е информационна наука и следователно за тази цел се нуждае от адекватни математически модели. Молекулярната биоматематика от своя страна се грижи за проектирането и създаването на тези модели. В основната част курса нашата главна задача ще бъде да разгледаме как по математически път можем да прогнозираме генетичната промяна на даден биологичен вид с течение на времето. За да можем да подготвим нашите математически пособия, видът и типът на информацията трябва да бъдат уточнени. Нека тогава припомним абстракцията на основните структурни нива на живата материя.



Така построено нашето дърво показва, че генетичните промени трябва да имат въздействие върху всички биологични нива нагоре, т.е. генетичните проучвания имат общо биологическо смисъл. От друга страна ходенето "нагоре" по дървото изисква все повече обобщения и затова значението на частичните данни се губи. По тази причина, за да имаме качествени обобщения трябва да обработим достатъчно данни, което пък изисква уеднаквен математически подход. Чрез следващата графика ще конкретизираме представянето на генетичната информация.



В клетките съществуват структури от ДНК наречени хромозоми. ДНК е полимер, изграден от множество нуклеотиди. Всеки нуклеотид се състои от нуклеотидна база, захар и една или няколко фосфатни групи. Има два

типа нуклеотидни бази - пурины и пиримидини. В ДНК има два вида пурины - аденин (A) и гуанин (G) и два вида пиримидини - тимин (T) и цитозин (C). Съчетанието от нуклеотидни бази, подредени една след друга върху нишката от ДНК, създава разнообразието от различни по вид и предназначение гени.

При определени условия гените се променят. Това се изразява като една нуклеотидна база преминава от един вид в друг вид, който е от същия или различен тип. Именно тази модификация представлява *мутацията*. Има два вида мутации - *транзиция* (от един вид в друг вид от същия тип) и *трансверсия* (от един вид във вид от другия тип). От математическа гледна точка това разграничаване има смисъл, тъй като трансверсии се наблюдават много по-рядко (например при човека шансовете да протече трансверсия са три пъти по-малки). Важно е да се отбележи, че в идеалистичния модел при мутация всяка променена база не изчезва, а само се променя. Следователно при мутация общият брой на нуклеотидите е константен. Към края на курса ще разгледаме и модели, при които на общия брой не е нужно да се гледа като на чисто константен, а по-скоро на константен, но с леко смущение, т.е. броят би могъл да се мени, но в някакви разумни (и очаквани) граници.

След като уточнихме понятията, за генетичната информация можем да кажем, че:

- има динамичен характер (т.е. подлежи на промяна (например чрез мутация) с течение на времето)
- четирите вида бази са нейните мономери (т.е. елементи) и техният брой в даден отрязък е константен (или поне остава приблизително еднакъв)

Естествено е тогава, моделът, който ще използваме, да бъде динамичен и да представя информацията в матричен вид. Това от своя страна ще доведе до необходимостта да работим на операторно ниво и да въведем някои закони, характерни за динамичните системи.

### Модел на Кимура

За да реализираме нашата цел, ще използваме един едновременно опростен но и елегантен подход, наречен модел на Кимура. Този модел трябва да бъде от операторен (матричен) тип и да включва в себе си вероятностите за транзиция и трансверсия. За експлицитното му представяне ще ползваме двупараметричния модел на Кимура (матрица, използваща модела на вероятностната РАМ матрица на Маргарет О. Дейхоф от 1979), а също и опростеният еднопараметричен модел на Джък-Кантор, при който шансовете за трансверсия и транзиция съвпадат, или предвид изключително малките шансове за настъпване на промяна в структурата на ДНК, ще разгледаме матрицата, като мутационна (т.е. единичната  $4 \times 4$  матрица, смутена с вероятностите за мутация). Пресмятането на тези матрици може да се разглежда в два аспекта – директно (директно пресмятане на стойностите), а също и като Марков процес (вероятностен процес, чието развитие във времето зависи също от текущото състояние, но не и от предишните; с други думи процес без "памет"). Засега ще обсъдим общи твърдения за модела, валидни и за двата варианта, а после ще преминем в детайл.

Нека с  $t \in [0, +\infty)$  отмерваме времето.

$n_i(t) \in \mathbb{N}$ , където  $i = \{A, G, C, T\}$  е броят на нуклеотидите с бази от тип  $i$  в даден отрязък ДНК.

Също така предвид изискванията за модела естествено е:

$$n_A(t) + n_C(t) + n_G(t) + n_T(t) = \text{const} \quad (\text{нарича се още закон за запазване структурата на живата материя})$$

Нека дефинираме векторът  $n(t)$ , такъв че

$$n(t) = (n_A(t), n_G(t), n_C(t), n_T(t))$$

Тогава моделът на Кимура гласи:

$$(1) \quad n(t) = M^t n(0)$$

където  $M$  е матрицата на линейния оператор  $M$ , който според конкретните вероятности за появата на промяна в броя на всеки една от базите изчислява броя им в следващ момент. Тази матрица ще наричаме *матрица на Кимура*, а конкретният ѝ вид зависи от нашата интерпретация. Представянето (1) не е необичайно. То следва от

$$(1.1) n(t+1) = M(n(t)) \Leftrightarrow n(t+1) = M n(t)$$

$$(1.2) M(M(M(\dots M(x)\dots))) = MMM\dots x = M^t x$$

По-общо казано равенството (1) за произволна матрица  $M$  всъщност е законът за еволюция при необратимост във времето. Още преди да въведем конкретния вид на  $M$  (която, разбира се, се очаква да бъде  $4 \times 4$  матрица) можем да направим лек коментар върху положителните и отрицателните аспекти на модела. Към положителната му страна трябва да припишем неговата елегантна, от математическа гледна точка, простота – на базата на експериментални данни за вероятностите на мутация само чрез пресмятането на степен на матрица можем да направим прогнози за състоянието на генетичната информация след определено време. По-нататък ние ще покажем, че този модел действа и за отрицателни стойности на времето, т.е. по данни за текущото състояние на даден вид можем да съдим за неговите генетични белези в миналото, което от исторически план несъмнено има стойност. Всичко това е добър предтекст за ползата от анализирането на модела. И все пак той има и своите недостатъци. За цялостта на генетичната информация се изисква да знаем не само количествата на нуклеотиди с определен вид бази, но и тяхната конкретна подредба. Например, нека за някакъв кодон (т.е. последователност от три нуклеотида, имаща генетичен смисъл) имаме 2 бази аденин и една цитозин, т.е. за някакво конкретно време  $t$  имаме  $n(t) = (2,0,1,0)$ . На този вектор обаче отговарят 3 кодона – AAC ACA CAA – това придава известна недетерминираност на прогнозите на модела.

#### Двупараметричен модел на Кимура.

Първото представяне на  $M$ , което ще разгледаме се нарича двупараметричен модел на Кимура (благодарение на двата параметъра присъстващи в матрицата  $M$ )

Нека  $\alpha$  е вероятността за транзичия, а  $\beta$  е вероятността за трансверсия. Като шансовете за протичане на мутация са изключително малки (трансверсия се наблюдава значително по рядко от транзичия), т.е.  $0 < \beta \ll \alpha \ll 1$ . Тогава:

$$(2) M = \begin{pmatrix} 1 - \alpha - 2\beta & \alpha & \beta & \beta \\ \alpha & 1 - \alpha - 2\beta & \beta & \beta \\ \beta & \beta & 1 - \alpha - 2\beta & \alpha \\ \beta & \beta & \alpha & 1 - \alpha - 2\beta \end{pmatrix}$$

Където  $\alpha$  и  $\beta$  са константи за всеки отделен вид.

#### Директен аспект на модела на Кимура

Да разгледаме модела директно означава да гледаме на  $M$  като на матрица, чийто елементи са функции на  $\alpha$  и  $\beta$  и чрез теорията на линейната алгебра да изведем своите заключения.

От курсовете по линейна алгебра знаем, че пресмятането на степен на матрица може да се улесни чрез пресмятането на която представлява жорданова матрица или в най – добрия случай – чисто диагонална матрица. Смятането на степен на :

$$J_{\text{cell}} = \begin{pmatrix} a & 1 & 0 & 0 & 0 \\ 0 & a & 1 & 0 & 0 \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & a & 1 \\ 0 & 0 & 0 & 0 & a \end{pmatrix} \text{ и нека } \text{size}(J_{\text{cell}}) = n \text{ е редът на } J_{\text{cell}}$$

С други думи матрицата с елементи  $a$  по диагонала и 1 под диагонала. Жорданова клетка от ред 1 наричаме тривиална жс

$$J = \begin{pmatrix} J_{\text{cell}_1} & 0 & 0 & 0 & 0 \\ 0 & J_{\text{cell}_2} & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & J_{\text{cell}_{k-1}} & 0 \\ 0 & 0 & 0 & 0 & J_{\text{cell}_k} \end{pmatrix}, \text{ като } m = \sum_{i=1}^k \text{size}(J_{\text{cell}_i})$$

Нека например имаме жордановите клетки

$$J_{\text{cell}_1} = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}, J_{\text{cell}_2} = \begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{pmatrix}, J_{\text{cell}_3} = (7), J_{\text{cell}_4} = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$$

Тогава жордановата матрица  $J$  съставена от тези клетки е от ред  $2 + 3 + 1 + 2 = 8$  и има вида

$$J = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 7 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{pmatrix}$$

Тогава диагонална матрица, т.е. матрица, чийто ненулеви елементи са само по диагонала, можем да наричаме жорданова матрица, съставена само от тривиални жорданови клетки. Сега да дефинираме степен на Жорданова клетка

$$J_{\text{cell}}^t = \begin{pmatrix} a_1^t & 1 & 0 & 0 & 0 \\ 0 & a_2^t & 1 & 0 & 0 \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & a_{n-1}^t & 1 \\ 0 & 0 & 0 & 0 & a_n^t \end{pmatrix}$$

Тогава за степента на жордановата матрица  $J$  имаме

$$J^t = \begin{pmatrix} J_{\text{cell}_1}^t & 0 & 0 & 0 & 0 \\ 0 & J_{\text{cell}_2}^t & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & J_{\text{cell}_{k-1}}^t & 0 \\ 0 & 0 & 0 & 0 & J_{\text{cell}_k}^t \end{pmatrix}$$

Теорията казва, че за каноничната форма  $N$  на произволна матрица  $M$  имаме

$$N = H^{-1} M H \Leftrightarrow M = H N H^{-1} \Rightarrow M^t = (H N H^{-1})^t = H N^t H^{-1},$$

където  $H$  е матрицата на прехода от  $M$  към  $N$ , а  $N$  е жорданова матрица. Тогава

нашата бъдеща задача става при конкретно представяне на  $M$  да намерим  $N$ ,  $N^t$ ,  $H$  и  $H^{-1}$ .

Нека припомним алгоритъма за намиране на каноничната форма  $N$  и матрицата на прехода  $H$  за матрицата  $M$ .

Както казахме,  $N$  е жорданова матрица, т.е. матрица с жорданови клетки по диагонала. Тогава трябва да уточним броя на клетките, елементите и редът на всяка клетка. Според теорията най-общо това, което трябва да направим е:

1. Намираме собствените стойности на  $M$
2. За всяка собствена стойност намираме съответната матрица  $B_i = M - \lambda_i I$
2. Изясняваме броя и реда на жордановите клетки чрез ранга на  $B_i$ .  $N$  се формира от тези жорданови клетки. (т.е. намираме  $N$ )
3. Намираме собствените (и присъединени вектори, ако има такива) на  $M$  чрез  $B_i$ .  $H$  се формира от тези вектори. (т.е. намираме  $H$ )
4. Намираме  $H^{-1}$

Ето как стоят нещата в детайл.

Първо намираме собствените стойности на матрицата  $M$ . Собствените стойности са решенията  $\lambda_i$  на уравнението  $\det(M - \lambda I) = 0$ . В нашия случай имаме

$$\det \begin{pmatrix} 1 - \alpha - 2\beta - \lambda & \alpha & \beta & \beta \\ \alpha & 1 - \alpha - 2\beta - \lambda & \beta & \beta \\ \beta & \beta & 1 - \alpha - 2\beta - \lambda & \alpha \\ \beta & \beta & \alpha & 1 - \alpha - 2\beta - \lambda \end{pmatrix} = 0$$

Как се смята детерминанта знаем – можем първо чрез елементарни преобразувания да занулим първите елементи на 2, 3 и 4 ред и после да развием по първия стълб. След това правим същото за  $3 \times 3$  матрицата, която остава и така докато получим уравнението. Така получаваме:

$$1 - 4x + 6x^2 - 4x^3 + x^4 - 4\alpha + 12x\alpha - 12x^2\alpha + 4x^3\alpha + 4\alpha^2 - \\ 8x\alpha^2 + 4x^2\alpha^2 - 8\beta + 24x\beta - 24x^2\beta + 8x^3\beta + 24\alpha\beta - 48x\alpha\beta + 24x^2\alpha\beta - \\ 16\alpha^2\beta + 16x\alpha^2\beta + 20\beta^2 - 40x\beta^2 + 20x^2\beta^2 - 32\alpha\beta^2 + 32x\alpha\beta^2 - 16\beta^3 + 16x\beta^3 = 0$$

След като опростим получаваме:

$$(x - 1)(x - 1 + 4\beta)(x - 1 + 2\alpha + 2\beta)^2 = 0$$

Т.е. за  $M$  имаме собствените стойности:

$$\lambda_1 = 1, \lambda_2 = 1 - 4\beta, \lambda_{3,4} = 1 - 2\alpha - 2\beta$$

Сега нека намерим  $B_i$

$$B_1 = \begin{pmatrix} 1 - \alpha - 2\beta & \alpha & \beta & \beta \\ \alpha & 1 - \alpha - 2\beta & \beta & \beta \\ \beta & \beta & 1 - \alpha - 2\beta & \alpha \\ \beta & \beta & \alpha & 1 - \alpha - 2\beta \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -\alpha - 2\beta & \alpha & \beta & \beta \\ \alpha & -\alpha - 2\beta & \beta & \beta \\ \beta & \beta & -\alpha - 2\beta & \alpha \\ \beta & \beta & \alpha & -\alpha - 2\beta \end{pmatrix}$$

$$\begin{aligned}
B_2 &= \begin{pmatrix} 1-\alpha-2\beta & \alpha & \beta & \beta \\ \alpha & 1-\alpha-2\beta & \beta & \beta \\ \beta & \beta & 1-\alpha-2\beta & \alpha \\ \beta & \beta & \alpha & 1-\alpha-2\beta \end{pmatrix} - \begin{pmatrix} 1-4\beta & 0 & 0 & 0 \\ 0 & 1-4\beta & 0 & 0 \\ 0 & 0 & 1-4\beta & 0 \\ 0 & 0 & 0 & 1-4\beta \end{pmatrix} = \\
& \begin{pmatrix} 2\beta-\alpha & \alpha & \beta & \beta \\ \alpha & 2\beta-\alpha & \beta & \beta \\ \beta & \beta & 2\beta-\alpha & \alpha \\ \beta & \beta & \alpha & 2\beta-\alpha \end{pmatrix} \\
B_{3,4} &= \begin{pmatrix} 1-\alpha-2\beta & \alpha & \beta & \beta \\ \alpha & 1-\alpha-2\beta & \beta & \beta \\ \beta & \beta & 1-\alpha-2\beta & \alpha \\ \beta & \beta & \alpha & 1-\alpha-2\beta \end{pmatrix} - \\
& \begin{pmatrix} 1-2\alpha-2\beta & 0 & 0 & 0 \\ 0 & 1-2\alpha-2\beta & 0 & 0 \\ 0 & 0 & 1-2\alpha-2\beta & 0 \\ 0 & 0 & 0 & 1-2\alpha-2\beta \end{pmatrix} = \begin{pmatrix} \alpha & \alpha & \beta & \beta \\ \alpha & \alpha & \beta & \beta \\ \beta & \beta & \alpha & \alpha \\ \beta & \beta & \alpha & \alpha \end{pmatrix}
\end{aligned}$$

За всяка  $k$ -кратна собствена стойност имаме  $f \leq k$  на брой жорданови клетки от ред не по-голям от  $k$ , като  $\sum_{i=1}^f \text{order}(\text{Jcell}_i) = k$ . Когато  $k = 1$ , тогава имаме тривиална жорданова клетка. Ако  $k > 1$ , то трябва да направим анализ.

В нашия случай имаме двете тривиални жорданови клетки  $(1)$ ,  $(1-4\beta)$ , съответстващи на собствените стойности  $\lambda_1$  и  $\lambda_2$ , а за  $\lambda_{3,4}$  трябва да помислим. Според теоремата за ранга, рангът на матрицата  $B_i$  е равен на броя на жордановите клетки, т.е.  $\text{rank}(B_i) = f$ . В нашия случай имаме  $\text{rank}(B_{3,4}) = 2$ , т.е. имаме 2 броя жорданови клетки, но тъй като те са от най-много втори ред, то следва, че имаме две тривиални жорданови клетки  $(1-2\alpha-2\beta)$ . Тъй като имаме само клетки от първи ред, то матрицата  $N$  следва да е диагонална. Сега можем да я построим. Поместваме клетките, така както ни харесва, т.е. без да спазваме определен ред. Тогава за  $N$  имаме

$$N = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-4\beta & 0 & 0 \\ 0 & 0 & 1-2\alpha-2\beta & 0 \\ 0 & 0 & 0 & 1-2\alpha-2\beta \end{pmatrix}$$

Сега за всяка собствена стойност трябва да намерим съответните собствен и присъединени (ако има такива) вектори. Както вече видяхме на всяка собствена стойност съответсват  $f$  на брой жорданови клетки. На всяка жорданова клетка от ред  $m$  съответстват 1 собствен вектор и  $m-1$  присъединени вектора. Собствените вектори  $v$  на собствената стойност  $\lambda_i$  намираме от формулата  $B_i v = 0$ , където  $B_i = M - \lambda_i I$ , а различните  $v$  (т.е. отговарящи на различна жорданова клетка за  $\lambda_i$ ) намираме като варираме свободните координати. Присъединени вектори  $v_{\text{assoc}j}$  намираме от формулата  $B_i v_1 = v$ ,  $B_i v_j = v_{j-1}$ , където отново може да се наложи да варираме по координати, ако имаме няколко

нетривиални жорданови клетки. След като сме намерили векторите, отговарящи на жордановите клетки,  $N$  се получава като запишем векторите един до друг, всяка групичка вектори (собствен + присъединени, като присъединените са по ред на получаване) отговаряща на съответната им жорданова клетка. В случая имаме само тривиални жорданови клетки и затова ще търсим само собствените вектори.

На  $\lambda_1$  отговаря  $v_1$ , такъв че при  $v_1 = (p \ q \ r \ s)^t$  имаме  $B_1 v_1 = 0$ , смятайки системата от 4 линейни уравнения виждаме, че  $v_1 = (p \ p \ p \ p)^t$ , тогава нека  $v_1 = (1 \ 1 \ 1 \ 1)^t$

На  $\lambda_2$  отговаря  $v_2$ , такъв че  $v_2 = (p \ q \ r \ s)^t$  имаме  $B_2 v_2 = 0$ , смятайки системата от 4 линейни уравнения виждаме, че  $v_2 = (p \ p \ -p \ -p)^t$ , тогава нека  $v_2 = (-1 \ -1 \ 1 \ 1)^t$

На  $\lambda_{3,4}$  отговаря  $v_{3,4}$ , такъв че  $v_{3,4} = (p \ q \ r \ s)^t$  имаме  $B_{3,4} v_{3,4} = 0$ , смятайки системата от 4 линейни уравнения виждаме, че  $v_{3,4} = (p \ -p \ r \ -r)^t$ , тогава нека  $v_3 = (0 \ 0 \ -1 \ 1)^t$ , а  $v_4 = (-1 \ 1 \ 0 \ 0)^t$

Имайки собствените вектори и предвид подредбата на жордановите клетки (съответстващи на собствените стойности) в каноничния вид (т.е.  $N$ ) имаме и матрицата на прехода  $H$ :

$$H = \begin{pmatrix} 1 & -1 & 0 & -1 \\ 1 & -1 & 0 & 1 \\ 1 & 1 & -1 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

Сега остава да намерим обратната на  $H$  матрица, т.е.  $H^{-1}$

Според теорията на линейната алгебра

$H^{-1} = \frac{1}{\det(H)} A_H^t$ , където  $A_H^t$  е транспонираната матрица на адюнгираната на  $H$  матрица, където

$$A_H = \{a_{ij} : a_{ij} = (-1)^{i+j} \Delta_{ij},$$

където  $\Delta_{ij}$  е поддетерминанта на  $H$ ,

такава че елементите от  $i$ -тия и  $j$ -тия ред/стълб не участват}

$$\text{Например } a_{11} = (-1)^{1+1} \det \begin{pmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix} = 2$$

Тъй като  $\det(H) = 8$ , тогава

$$H^{-1} = \frac{1}{8} \begin{pmatrix} 2 & -2 & 0 & -4 \\ 2 & -2 & 0 & 4 \\ 2 & 2 & -4 & 0 \\ 2 & 2 & 4 & 0 \end{pmatrix}^t = \frac{1}{4} \begin{pmatrix} 1 & -1 & 0 & -1 \\ 1 & -1 & 0 & 1 \\ 1 & 1 & -1 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}^t = \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 0 & 0 & -2 & 2 \\ -2 & 2 & 0 & 0 \end{pmatrix}$$

Тогава за  $M^t$  имаме

$$M^t = \begin{pmatrix} 1 & -1 & 0 & -1 \\ 1 & -1 & 0 & 1 \\ 1 & 1 & -1 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & (1-4\beta)^t & 0 & 0 \\ 0 & 0 & (1-2\alpha-2\beta)^t & 0 \\ 0 & 0 & 0 & (1-2\alpha-2\beta)^t \end{pmatrix} \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 0 & 0 & -2 & 2 \\ -2 & 2 & 0 & 0 \end{pmatrix}$$

Замествайки собствените стойности в  $N$  с техните имена получаваме

$$\begin{aligned}
M^t &= \frac{1}{4} \begin{pmatrix} 1 & -1 & 0 & -1 \\ 1 & -1 & 0 & 1 \\ 1 & 1 & -1 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda_2^t & 0 & 0 \\ 0 & 0 & \lambda_{3,4}^t & 0 \\ 0 & 0 & 0 & \lambda_{3,4}^t \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 0 & 0 & -2 & 2 \\ -2 & 2 & 0 & 0 \end{pmatrix} = \\
&= \frac{1}{4} \begin{pmatrix} 1 & -\lambda_2^t & 0 & -\lambda_{3,4}^t \\ 1 & -\lambda_2^t & 0 & \lambda_{3,4}^t \\ 1 & \lambda_2^t & -\lambda_{3,4}^t & 0 \\ 1 & \lambda_2^t & \lambda_{3,4}^t & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 0 & 0 & -2 & 2 \\ -2 & 2 & 0 & 0 \end{pmatrix} = \\
&= \frac{1}{4} \begin{pmatrix} 1 + \lambda_2^t + 2\lambda_{3,4}^t & 1 + \lambda_2^t - 2\lambda_{3,4}^t & 1 - \lambda_2^t & 1 - \lambda_2^t \\ 1 + \lambda_2^t - 2\lambda_{3,4}^t & 1 + \lambda_2^t + 2\lambda_{3,4}^t & 1 - \lambda_2^t & 1 - \lambda_2^t \\ 1 - \lambda_2^t & 1 - \lambda_2^t & 1 + \lambda_2^t + 2\lambda_{3,4}^t & 1 + \lambda_2^t - 2\lambda_{3,4}^t \\ 1 - \lambda_2^t & 1 - \lambda_2^t & 1 + \lambda_2^t - 2\lambda_{3,4}^t & 1 + \lambda_2^t + 2\lambda_{3,4}^t \end{pmatrix}
\end{aligned}$$

Тъй като знаем, че  $0 < \beta \ll \alpha \ll 1$ , то следователно и:

$$0 < \lambda_2 = 1 - 4\beta < 1$$

$$0 < \lambda_{3,4} = 1 - 2\alpha - 2\beta < 1$$

Тогава имаме още, че:

$$\lambda_2 \xrightarrow{t \rightarrow \infty} 0, \lambda_{3,4} \xrightarrow{t \rightarrow \infty} 0 \text{ и тогава:}$$

$$M^t \xrightarrow{t \rightarrow \infty} \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \text{ следователно } n(t) \xrightarrow{t \rightarrow \infty} \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} n(0)$$

Нека  $n(0) = (n_A, n_G, n_C, n_T)^t$ , и  $N = n_A + n_G + n_C + n_T$  т.е.

$$n(t) \xrightarrow{t \rightarrow \infty} \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} n_A \\ n_G \\ n_C \\ n_T \end{pmatrix}$$

умножавайки двете матрици в дясно достигаме до резултата



$$n(t) \xrightarrow{t \rightarrow \infty} \frac{1}{4} \begin{pmatrix} n_A + n_G + n_C + n_T \\ n_A + n_G + n_C + n_T \\ n_A + n_G + n_C + n_T \\ n_A + n_G + n_C + n_T \end{pmatrix} = \begin{pmatrix} \frac{N}{4} \\ \frac{N}{4} \\ \frac{N}{4} \\ \frac{N}{4} \end{pmatrix}$$

Тълкуването на резултата е ясно – с времето моделът показва тенденцията дяловото участие на нуклеотидните бази да се изравнява, т.е. независимо от началното съотношение на нуклеотидните бази цялостната тенденция на модела е мутацията върви към равновесното положение  $\left(\frac{N}{4} \quad \frac{N}{4} \quad \frac{N}{4} \quad \frac{N}{4}\right)$

### Матрицата на Кимура като мутационна матрица

Нека  $\alpha$  е вероятността за транзичия, а  $\beta$  е вероятността за трансверсия. Тъй като шансовете за протичане на мутация са изключително малки (трансверсия се наблюдава значително по рядко от транзичия), т.е.  $0 < \beta \ll \alpha \ll 1$ , то едно удобно представяне на матрицата  $M$  е като леко смущение на идентитета. Тогава:

$$M = I + \xi(\alpha, \beta)A \\ M^t = (I + \xi(\alpha, \beta)A)^t$$

Където  $\alpha$  и  $\beta$  са константи за всеки отделен вид,  $\xi(\alpha, \beta)$  е смущението,  $A$  е 4x4 матрица.

Предвид видът на  $M$  в дупараметричния модел на Кимура, който вече представихме, всъщност имаме

$$M - I = \xi(\alpha, \beta)A$$

Нека видим на какво е равна  $M - I$

$$M - I = \begin{pmatrix} 1 - \alpha - 2\beta & \alpha & \beta & \beta \\ \alpha & 1 - \alpha - 2\beta & \beta & \beta \\ \beta & \beta & 1 - \alpha - 2\beta & \alpha \\ \beta & \beta & \alpha & 1 - \alpha - 2\beta \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \\ = \begin{pmatrix} -\alpha - 2\beta & \alpha & \beta & \beta \\ \alpha & -\alpha - 2\beta & \beta & \beta \\ \beta & \beta & -\alpha - 2\beta & \alpha \\ \beta & \beta & \alpha & -\alpha - 2\beta \end{pmatrix} = \\ = (\alpha + 2\beta) \begin{pmatrix} -1 & \frac{\alpha}{\alpha + 2\beta} & \frac{\beta}{\alpha + 2\beta} & \frac{\beta}{\alpha + 2\beta} \\ \frac{\alpha}{\alpha + 2\beta} & -1 & \frac{\beta}{\alpha + 2\beta} & \frac{\beta}{\alpha + 2\beta} \\ \frac{\beta}{\alpha + 2\beta} & \frac{\beta}{\alpha + 2\beta} & -1 & \frac{\alpha}{\alpha + 2\beta} \\ \frac{\beta}{\alpha + 2\beta} & \frac{\beta}{\alpha + 2\beta} & \frac{\alpha}{\alpha + 2\beta} & -1 \end{pmatrix}$$

Ако положим  $\gamma = \frac{\alpha}{2\beta}$ , което всъщност е отношението транзичия/трансверсия, то получаваме

$$\frac{\alpha + 2\beta}{\alpha} = 1 + \frac{1}{r} = \frac{r+1}{r}; \quad \frac{\alpha + 2\beta}{\beta} = \frac{\alpha}{\beta} + 2 = 2r + 2 = 2(r+1); \quad \alpha + 2\beta = 2\beta \frac{\alpha + 2\beta}{2\beta} = 2\beta(r+1)$$

Тогава за  $M - I$  имаме:

$$M - I = 2\beta(r+1) \begin{pmatrix} -1 & \frac{r}{r+1} & \frac{1}{2(r+1)} & \frac{1}{2(r+1)} \\ \frac{r}{r+1} & -1 & \frac{1}{2(r+1)} & \frac{1}{2(r+1)} \\ \frac{1}{2(r+1)} & \frac{1}{2(r+1)} & -1 & \frac{r}{r+1} \\ \frac{1}{2(r+1)} & \frac{1}{2(r+1)} & \frac{r}{r+1} & -1 \end{pmatrix} = 2\beta(r+1) A(r), \text{ където } A(r) \text{ е последната матрица}$$

Окончателно:

$$M = I + 2\beta(1+r) A(r) \\ M^t = (I + 2\beta(1+r) A(r))^t$$

### Коментар

Интересно е как  $A(r)$  стана вече еднопараметричен оператор. Това ще даде доста гъвкавост на модела в по-нататъшните математически и биологически интерпретации. Предвид ролята си  $A$  ще наричаме *моментна мутационна матрица*.

### Молекулярната еволюция като процес на Марков

Нека изясним понятията като се запознаем с характеристиките на Марковите процеси.

Нека имаме система от обекти, които са в дадено, характерно за тях състояние (т.е. състояние което е допустимо да приемат) и в този смисъл обобщавайки, казваме че системата е в някакво дадено състояние. С течение на времето по явни или не явни правила тези обекти се променят и системата преминава от едно състояние в друго. Именно тази характеристика на промяна на системата ще наричаме *процес*. Последователността от състояния на процеса се нарича *верига* от състоянията на процеса. Процесите бихме могли да асоциираме с техните вериги. Естественият математически модел, пълноценно илюстриращ верига на процес е редицата. Процесите могат да бъдат разглеждани като *непрекъснати* и *дискретни* в зависимост в зависимост от това дали редицата представяща техните състояния има непрекъснат или дискретен характер.

Марковите процеси от своя страна са процеси, при които бъдещото състояние на системата не зависи от предишните състояния, а зависи единствено от настоящето състояние. С други думи, описанието на текущото състояние включва всичката информация, която може да повлияе на бъдещото развитие на процеса. Заради това си свойство тези процеси често биват наричани процеси без памет.

Абстрактната природа на Марковите процеси е причина те да способстват описването (въпреки понякога в опростен план) на процеси от най-различен характер – във физиката, статистиката, модели от теорията на игрите, биоматематиката и др. науки. Въпреки абстрактното си естество, Марковите процеси най-често описват вероятностни процеси. Във всяко състояние на процеса се разглежда система от вероятности, зависещи от текущото състояние, на базата на които се извършва прехода към следващото състояние.

Идеята за отгъждествяване на молекулярната еволюция (в нашия случай ние разглеждаме матрицата на Кимура еволюираща във времето) с процес на Марков произлиза от естествената необходимост за абстракция на математическите модели. Твърде конкретните модели имат недостатъка често да дават нереалистични прогнози, поради ограничеността на своето приложение. В досегашните интерпретации на модела на Кимура, разглеждахме параметрите  $\alpha$  и  $\beta$  (транзиция / трансверсия) като константни по време на еволюционния процес на генетичната информация. Предвид факта обаче, че тези параметри в реалния случай зависят от средата, в която даден вид е разглеждан, а с времето средата се променя (т.е. средата зависи от времето), един по-приемлив модел би бил, ако параметрите се разглеждаха нека като константни, а като зависими от времето. Взимайки предвид това ново изискване, теорията на Марковите процеси би спомогнала за осъществяването на модела в този му нов аспект.

Сега казаното за Марковите модели ще го конкретизираме с пособията на математиката. Както казахме по-рано, процесите ще характеризираме с тяхните вериги.

Нека  $S = (e_1, e_2, \dots, e_n)$  е материална система от  $n$  на брой елемента. Всеки нейн елемент може да заема  $m$  на брой състояния. Нека тези елементи с времето се прегрупират в  $m$  на брой подсистеми, според възможните си състояния. Да означим с  $x_1, x_2, \dots, x_m$  количествените мерки на съответните подсистеми.

Нека например имаме системата  $S = (a, b)$ , чиито елеченти могат да приемат състоянията  $A, B, C$ , тогава за  $S$  имаме следните подсистеми:

$$S_A = (x), S_B = (x), S_C = (x), x \in \{a, b\}, \text{ където } x \text{ е съответно в състояние } A, B \text{ и } C$$

Ще казваме, че имаме непрекъсната Маркова верига за системата  $S$ , ако са изпълнени следните три изисквания: За една  $m$ -торка величини  $x_1, \dots, x_m$ , биващи количествени мерки на подсистемите на една материална система  $S$ , ще казваме, че образуват непрекъсната Маркова верига, ако за тези величини са в сила следните условия:

(i) Величините  $x_j = x_j(t)$  са непрекъснати за  $t \geq 0, j = 1, \dots, m$  и е изпълнено следното

$$(*) \sum_{j=1}^m x_j(t) = \text{const за } t \geq 0$$

Където  $(*)$  се нарича закон за запазване на системата. Тъй като в случая на Кимура това беше закона за запазване на структурата на живата материя, който вече разгледахме в нашите съждения, интерпретацията на  $(*)$  тук е ясна – сумата на количествените мярки остава константна при трансформации на системата.

(ii) Възможни са преходи от дадена  $i$ -та подсистема към друга  $j$ -та подсистема с вероятност:

$0 \leq p_{ij} \leq 1$ , като  $p_{ij} = p_{ij}(t)$  – непрекъснати при  $t \geq 0$  и е в сила

$$(**.1) \sum_{j=1}^m p_{ij}(t) = 1 \text{ за } t \geq 0, i = 1, \dots, m$$

Анализирайки (ii) стигаме до извода, че количествената мярка на  $i$ -тата подсистема  $x_i(t)$  се определя спрямо мярките  $x_1^0, x_2^0, \dots, x_m^0$  в началния момент  $t = 0$  по следния принцип:

(1) от остатъка  $p_{ii}(t) x_i^0$  или с други думи прихода на елементи от една система в същата

(2) от прихода на елементи от останалите подсистеми

Т.е.  $x_i(t)$  се определя от прихода на елементи от всички подсистеми, включително и  $i$ :

$$(**.2) x_i(t) = \sum_{j=1}^m p_{ij}(t) x_j^0 \text{ за всяко } i = 1, \dots, m$$

Нека  $x(t) = (x_1(t), \dots, x_m(t))$ ,  $x^0 = (x_1^0, \dots, x_m^0)$ , където  $x^0$  са количествените стойности за началния момент  $t = 0$  и следователно естествено е да се очаква  $x^0 = x(0)$ , а  $P(t)$  е матрична функция съпоставяща на всяко  $t \geq 0$  матрица на вероятностите за всевъзможните преходи, т.е. за  $P(t)$  имаме:

$$P(t) = \{p_{ij}(t)\}_{m \times m} = \begin{pmatrix} p_{11}(t) & \cdots & p_{1m}(t) \\ \vdots & \ddots & \vdots \\ p_{m1}(t) & \cdots & p_{mm}(t) \end{pmatrix}$$

Предвид (i) и (ii) в сила е следното представяне.

$$(\#) x(t) = P(t) x_0$$

( # ) се нарича формула за общото решение или конкретно в нашия случай – формула за еволюирания вектор на състоянията.  $P(t)$  ще наричаме матрица на прехода. С въпроса дали в действителност съществуват обекти от вида на  $P(t)$  и тяхното конкретно описание ще се заемем по-нататък.

(iii) Еднозначност на процеса

След като изяснихме смисъла на  $P(t)$  се налага с оглед еднозначност на протичането на процеса, т.е. процеса да еволюира линейно (което пък е необходимо за съществуването на еднозначност на веригата на процеса), въвеждането на следните изисквания:

$$\begin{aligned} (1) P(0) &= I \\ (2) P(t+s) &= P(t)P(s) \quad \forall t, s \geq 0 \\ (3) P(t+s) &= P(s+t), \text{ т.е.} \\ &P(s)P(t) = P(t)P(s) \end{aligned}$$

Нека сега разгълкуваме горната система от изисквания.

(1) Възниква естествено от изискването  $x(0) = x_0$ , тъй като  $x(0) = P(t)x^0 \Leftrightarrow x(0) = P(t)x(0) \Leftrightarrow P(t) = I$

(2) Нарича се *равенство на Чапман-Колмогоров*, което е и основен принцип в динамичните системи (каквато е и нашата). Това значи, че системата ще еволюира еднакво за време  $t+s$  започвайки от 0 и при интервала  $s$  с начално  $t$ .

(3) Условието придава единственост (линейност) на еволюционния процес – независимо от пътя на еволюцията (т.е. състоянията през които минава системата) за еднакъв интервал от време и еднакви начални количествени мерки системата еволюира до едно и също свое състояние.

Възниква въпросът съществува ли изобщо и от какъв тип е матрицата  $P$ . За да отговорим на този въпрос, ще се наложи да проведем анализ.

#### Диференцируеми операторни функции

Нека  $\{P(t), t \geq 0\}$  е 1-параметрична фамилия от линейни оператори, т.е. дадена е операторната функция

$$P(t) : \mathbb{R}^m \rightarrow \mathbb{R}^m, t \geq 0$$

Ще казваме, че съществува  $\dot{P}(t_0)$  или че  $P(t)$  е диференцируема в  $t = t_0$ , ако съществува линейният оператор

$$\begin{aligned} Q(t_0) : \mathbb{R}^m &\rightarrow \mathbb{R}^m, \\ \lim_{\theta \rightarrow 0} \frac{P(t_0 + \theta) - P(t_0)}{\theta} &= Q(t_0) \end{aligned}$$

Тогава границата  $Q(t_0)$  на това диференчно частно означаваме с  $\dot{P}(t_0)$

Ако приложим тази дефиниция към 1-параметрична фамилия от линейни оператори, за които са изпълнени (iii.2) и (iii.3), т.е.

$$\begin{aligned} (iii.2) P(0) &= I \\ (iii.3) P(t+s) &= P(t)P(s) \end{aligned}$$

То ще получим:

$$(\blacklozenge) \frac{P(t+\theta) - P(t)}{\theta} \stackrel{(iii.3)}{=} \frac{P(t) + P(\theta) - P(t)}{\theta} = \frac{P(\theta) - I}{\theta} P(t) \stackrel{(iii.2)}{=} \frac{P(\theta) - P(0)}{\theta} P(t)$$

Следователно

$$\begin{aligned} \dot{P}(t) &\stackrel{\text{def}}{=} \lim_{\theta \rightarrow 0} \frac{P(t+\theta) - P(t)}{\theta} = \lim_{\theta \rightarrow 0} \left( \frac{P(\theta) - P(0)}{\theta} P(t) \right) \stackrel{(\clubsuit)}{=} \lim_{\theta \rightarrow 0} \left( \frac{P(\theta) - P(0)}{\theta} \right) \lim_{\theta \rightarrow 0} (P(t)) = \\ &= \lim_{\theta \rightarrow 0} \left( \frac{P(0+\theta) - P(0)}{\theta} \right) P(t) \stackrel{\text{def}}{=} \dot{P}(0) \cdot P(t), \text{ т.е. получихме:} \end{aligned}$$

$$\dot{P}(t) = \dot{P}(0) P(t)$$

Т.е. с други думи получихме, че 1-параметричната фамилия от линейни оператори, за които са изпълнени (iii.2) и (iii.3), е диференцируема  $\forall t \Leftrightarrow$  е диференцируема за  $t = 0$  или с други думи  $\exists \dot{P}(t) \Leftrightarrow \exists \dot{P}(0)$

Преди да продължим напред, нека направим лек коментар върху (iii). Еднопараметрична фамилия от оператори  $P(t)$ , такава че:

$$\{P(t) : \mathbb{R}^m \rightarrow \mathbb{R}^m, \forall t \geq 0 \mid P(0) = I, P(t+s) = P(t)P(s), P(t+s) = P(s+t), \forall t, s \geq 0\}$$

ще наричаме еднопараметрична полугрупа от линейни оператори от  $\mathbb{R}^m$  в  $\mathbb{R}^m$ , тъй като всички оператори от вида на  $P(t)$  всъщност образуват полугрупа. Ако пък имаме оператори от вида на  $P(t)$ , но при условие че  $t, s \in \mathbb{R}^1$ , то вече те ще образуват еднопараметрична група от линейни оператори.

Нека положим  $A = \dot{P}(0)$ . Тогава ще имаме

$$(\clubsuit) \dot{P}(t) = A P(t)$$

Т.е. получихме 1-параметричната полугрупа от линейни оператори:

$$(\star) \left\{ P(t) : \mathbb{R}^m \rightarrow \mathbb{R}^m \forall t \geq 0 \mid \dot{P}(t) = A P(t), P(0) = I \right\}$$

Предвид проведеня дотук анализ можем да отъждествим матриците  $P(t)$  на линейните оператори от полугрупата  $(\star)$  с матриците на прехода на всички Маркови вериги. Следователно, намирайки решението на задачата на Коши от операторен вид  $(\clubsuit)$  ще намерим вида на матриците на прехода на всички Маркови вериги, а с това, тъй като те ги характеризират изцяло, всъщност ще сме намерили вида на всички Маркови вериги.

Матрицата  $A$  на линейния оператор  $A$  ще наричаме генератор на Марковата верига.  $A$  в действителност представлява матрицата на скоростите на преходите малко след (в непосредствена близост до) момента  $t = 0$ . С други думи:

$$(\dagger) P(\theta) = I + \theta A, \theta \approx 0$$

Т.е.  $P(\theta)$  е малка деформация на единичната матрица  $I$ . Валидността на този факт е тривиална и се вижда при:

$$A = \dot{P}(0) = \lim_{\theta \rightarrow 0} \frac{P(0+\theta) - P(0)}{\theta} = \lim_{\theta \rightarrow 0} \frac{P(\theta) - I}{\theta} = \lim_{\theta \rightarrow 0} \frac{I + \theta A - I}{\theta} = \lim_{\theta \rightarrow 0} A = A$$

#### Решение на операторното ОДУ от първи ред

Решаването на  $(\clubsuit)$  се изисква първо операторното ОДУ от първи ред да се сведе до векторно ОДУ, което вече се свежда до решаване на система от ОДУ от първа степен. Практическият процес по извеждане на решението на това ОДУ тук няма да правя, тъй като доказателството не е тривиално и би обременило читателя изисквайки не малко концентрация и математически опит. Съответната теория е развита в книгите по ОДУ, като възнамерявам в следващо издание на материала да представя приложение върху проблема.

#### Теорема на Поанкаре

Нека имаме 1-параметричната група от линейни оператори:

$$(\star\star) \left\{ P(t) : \mathbb{R}^m \rightarrow \mathbb{R}^m \forall t \mid \dot{P}(t) = A P(t), P(0) = I, A = \dot{P}(0) \right\}$$

Когато  $(\star\star)$  е диференцируема, а  $A$  е диагонализуем оператор, то  $P(t)$  е от експоненциален тип или по-точно:

$$(\star\star\star) P(t) = e^{tA}, A = \dot{P}(0)$$

### Моделът на Кимура като модел на молекулярната еволюция по Марков

Сега вече бихме могли да преминем към практическото разглеждане на модела на Кимура като модел, описващ Марков процес, и да видим резултатите от въвеждането на този му аспект.

Първо обаче, трябва формално да определим, дали моделът на Кимура наистина описва процес на Марков – т.е. да видим какво говорят фактите, преди да се впускате в интерпретации. Нека си припомним модела на Кимура:

$n_i(t) \in \mathbb{N}$ , където  $i = \{A, G, C, T\}$  е броят на нуклеотидите с бази от тип  $i$  в даден отрязък ДНК.

$n_A(t) + n_C(t) + n_G(t) + n_T(t) = N = \text{const}$  (закон за запазване структурата на живата материя)

$n(t) = (n_A(t), n_G(t), n_C(t), n_T(t))$

А матрицата на Кимура  $M$  повдигната на степен  $t$  имаше вида:

$$M^t = \{m_{ij}\}_{4 \times 4} = \frac{1}{4} \begin{pmatrix} 1 + \lambda_2^t + 2\lambda_{3,4}^t & 1 + \lambda_2^t - 2\lambda_{3,4}^t & 1 - \lambda_2^t & 1 - \lambda_2^t \\ 1 + \lambda_2^t - 2\lambda_{3,4}^t & 1 + \lambda_2^t + 2\lambda_{3,4}^t & 1 - \lambda_2^t & 1 - \lambda_2^t \\ 1 - \lambda_2^t & 1 - \lambda_2^t & 1 + \lambda_2^t + 2\lambda_{3,4}^t & 1 + \lambda_2^t - 2\lambda_{3,4}^t \\ 1 - \lambda_2^t & 1 - \lambda_2^t & 1 + \lambda_2^t - 2\lambda_{3,4}^t & 1 + \lambda_2^t + 2\lambda_{3,4}^t \end{pmatrix}$$

Естествено е, предвид вероятностния си характер, матрицата на Кимура  $M^t$  да разглеждаме като матрица на преходите на Марковата верига. Т.е. имаме:

$$P(t) = M^t$$

За модела на Кимура се очаква да са в сила:

$$(i) \sum_{i=\{A,G,C,T\}} n_i(t) = N = \text{const} \forall t$$

$$(ii) \sum_{j=1}^4 m_{ij} = 1 \forall i = \{1, 2, 3, 4\}$$

$$(iii.1) P(0) = M^t|_{t=0} = M^0 = I$$

$$(iii.2) P(t+s) = P(t)P(s) = M^{t+s} = M^t M^s$$

$$(iii.3) P(t+s) = P(s+t) = M^{t+s} = M^{s+t}$$

Проверка на (i)

Поради закона за запазване на структурата на живата материя (i) е изпълнено по условие.

Проверка на (ii)

$$i = 1$$

$$\sum_{j=1}^4 m_{1j} = \frac{1}{4} (1 + \lambda_2^t + 2\lambda_{3,4}^t + 1 + \lambda_2^t - 2\lambda_{3,4}^t + 1 - \lambda_2^t + 1 - \lambda_2^t) = 1$$

Поради очевидното присъствие на едни и същи елементи на всеки ред в едни и същи количества, сумите на 4-те реда са равни. Предвид това и предвид успешната проверка, която проведехме, то следователно условие (ii) е изпълнено.

Проверка на (iii.1)

Тъй като  $P(0) = M^0$ , то условието очевидно е изпълнено, но за цялост:

$$\begin{aligned}
M^0 &= \frac{1}{4} \begin{pmatrix} 1 + \lambda_2^0 + 2\lambda_{3,4}^0 & 1 + \lambda_2^0 - 2\lambda_{3,4}^0 & 1 - \lambda_2^0 & 1 - \lambda_2^0 \\ 1 + \lambda_2^0 - 2\lambda_{3,4}^0 & 1 + \lambda_2^0 + 2\lambda_{3,4}^0 & 1 - \lambda_2^0 & 1 - \lambda_2^0 \\ 1 - \lambda_2^0 & 1 - \lambda_2^0 & 1 + \lambda_2^0 + 2\lambda_{3,4}^0 & 1 + \lambda_2^0 - 2\lambda_{3,4}^0 \\ 1 - \lambda_2^0 & 1 - \lambda_2^0 & 1 + \lambda_2^0 - 2\lambda_{3,4}^0 & 1 + \lambda_2^0 + 2\lambda_{3,4}^0 \end{pmatrix} = \\
&= \frac{1}{4} \begin{pmatrix} 1+1+2 & 1+1-2 & 1-1 & 1-1 \\ 1+1-2 & 1+1+2 & 1-1 & 1-1 \\ 1-1 & 1-1 & 1+1+2 & 1+1-2 \\ 1-1 & 1-1 & 1+1-2 & 1+1+2 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \frac{1}{4} 4I = I
\end{aligned}$$

Проверка на (iii.2)

$$M^{t+s} \stackrel{?}{=} M^t M^s$$

$$\begin{aligned}
M^t M^s &= \frac{1}{4} \begin{pmatrix} 1 + \lambda_2^t + 2\lambda_{3,4}^t & 1 + \lambda_2^t - 2\lambda_{3,4}^t & 1 - \lambda_2^t & 1 - \lambda_2^t \\ 1 + \lambda_2^t - 2\lambda_{3,4}^t & 1 + \lambda_2^t + 2\lambda_{3,4}^t & 1 - \lambda_2^t & 1 - \lambda_2^t \\ 1 - \lambda_2^t & 1 - \lambda_2^t & 1 + \lambda_2^t + 2\lambda_{3,4}^t & 1 + \lambda_2^t - 2\lambda_{3,4}^t \\ 1 - \lambda_2^t & 1 - \lambda_2^t & 1 + \lambda_2^t - 2\lambda_{3,4}^t & 1 + \lambda_2^t + 2\lambda_{3,4}^t \end{pmatrix} \\
&= \frac{1}{4} \begin{pmatrix} 1 + \lambda_2^s + 2\lambda_{3,4}^s & 1 + \lambda_2^s - 2\lambda_{3,4}^s & 1 - \lambda_2^s & 1 - \lambda_2^s \\ 1 + \lambda_2^s - 2\lambda_{3,4}^s & 1 + \lambda_2^s + 2\lambda_{3,4}^s & 1 - \lambda_2^s & 1 - \lambda_2^s \\ 1 - \lambda_2^s & 1 - \lambda_2^s & 1 + \lambda_2^s + 2\lambda_{3,4}^s & 1 + \lambda_2^s - 2\lambda_{3,4}^s \\ 1 - \lambda_2^s & 1 - \lambda_2^s & 1 + \lambda_2^s - 2\lambda_{3,4}^s & 1 + \lambda_2^s + 2\lambda_{3,4}^s \end{pmatrix} = \\
&= \frac{1}{16} \begin{pmatrix} 4 + 4\lambda_2^{s+t} + 8\lambda_{3,4}^{s+t} & 4 + 4\lambda_2^{s+t} - 8\lambda_{3,4}^{s+t} & 4 - 4\lambda_2^{s+t} & 4 - 4\lambda_2^{s+t} \\ 4 + 4\lambda_2^{s+t} - 8\lambda_{3,4}^{s+t} & 4 + 4\lambda_2^{s+t} + 8\lambda_{3,4}^{s+t} & 4 - 4\lambda_2^{s+t} & 4 - 4\lambda_2^{s+t} \\ 4 - 4\lambda_2^{s+t} & 4 - 4\lambda_2^{s+t} & 4 + 4\lambda_2^{s+t} + 8\lambda_{3,4}^{s+t} & 4 + 4\lambda_2^{s+t} - 8\lambda_{3,4}^{s+t} \\ 4 - 4\lambda_2^{s+t} & 4 - 4\lambda_2^{s+t} & 4 + 4\lambda_2^{s+t} - 8\lambda_{3,4}^{s+t} & 4 + 4\lambda_2^{s+t} + 8\lambda_{3,4}^{s+t} \end{pmatrix} = \frac{1}{4} M^{t+s}
\end{aligned}$$

С това проверката на (iii.2) е завършена успешно. Проверка на (iii.3) е тривиална, предвид вида на елементите на  $M^{t+s}$

В нашия случай, предвид теоремата на Поанкаре, имаме:

$$P(t) = e^{tA_0}, \text{ където } A_0 = \left. \frac{d}{dt} (M^t) \right|_{t=0}$$

Където  $A_0$  ще наричаме точен генератор на Марковата верига

Тъй като всеки от елементите на  $M^t$  е сума от константа и две показателни функции от вида  $a^t$  и тъй като  $\frac{d}{dt} (a^t) =$

$a^t \ln(a)$ , но при  $t = 0$  имаме  $\left. \frac{d}{dt} (a^t) \right|_{t=0} = a^0 \ln(a) = \ln(a)$ , то  $A_0$  съществува. Тогава за  $A_0$  имаме :

$$\begin{aligned}
\mathbf{A}_0 &= \frac{d}{dt} (\mathbf{M}^t) \Big|_{t=0} = \frac{1}{4} \begin{pmatrix} \ln \lambda_2 + 2 \ln \lambda_{3,4} & \ln \lambda_2 - 2 \ln \lambda_{3,4} & -\ln \lambda_2 & -\ln \lambda_2 \\ \ln \lambda_2 - 2 \ln \lambda_{3,4} & \ln \lambda_2 + 2 \ln \lambda_{3,4} & -\ln \lambda_2 & -\ln \lambda_2 \\ -\ln \lambda_2 & -\ln \lambda_2 & \ln \lambda_2 + 2 \ln \lambda_{3,4} & \ln \lambda_2 - 2 \ln \lambda_{3,4} \\ -\ln \lambda_2 & -\ln \lambda_2 & \ln \lambda_2 - 2 \ln \lambda_{3,4} & \ln \lambda_2 + 2 \ln \lambda_{3,4} \end{pmatrix} = \\
&= \frac{1}{4} \begin{pmatrix} \ln \lambda_2 \lambda_{3,4}^2 & \ln \frac{\lambda_2}{\lambda_{3,4}^2} & \ln \frac{1}{\lambda_2} & \ln \frac{1}{\lambda_2} \\ \ln \frac{\lambda_2}{\lambda_{3,4}^2} & \ln \lambda_2 \lambda_{3,4}^2 & \ln \frac{1}{\lambda_2} & \ln \frac{1}{\lambda_2} \\ \ln \frac{1}{\lambda_2} & \ln \frac{1}{\lambda_2} & \ln \lambda_2 \lambda_{3,4}^2 & \ln \frac{\lambda_2}{\lambda_{3,4}^2} \\ \ln \frac{1}{\lambda_2} & \ln \frac{1}{\lambda_2} & \ln \frac{\lambda_2}{\lambda_{3,4}^2} & \ln \lambda_2 \lambda_{3,4}^2 \end{pmatrix}
\end{aligned}$$