Міністрерство освіти і науки України

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «КИЄВО-МОГИЛЯНСЬКА АКАДЕМІЯ»

Кафедра математики факультету інформатики

МОДЕЛЮВАННЯ СОЦІАЛЬНИХ ПРОЦЕСІВ ЯК МАРКОВСЬКИХ ПРОЦЕСІВ ПРИЙНЯТТЯ РІШЕНЬ

Текстова частина до курсової роботи за спеціальністю «Прикладна математика»

Керівник курсової роботи

к.ф.-м.н., доц.____________

____

«____»___________2020 р.

Виконав студент___________

____

«____»___________2020 р.

Київ 2020

Зміст

ВСТУП

Тема роботи

Моделювання соціальних процесів як марковських процесів прийняття рішень.

Мета роботи

Розглянути марковські процеси та марковські поля, їх властивості, їх застосування в соціології та розглянути відповідні задачі.

Актуальність теми

Одним з найважливіших факторів в процесі прийняття рішень є фактор випадковості. При цьому варто зазначити, що так званий факт невизначеності не є еквівалентним фактору випадковості, так як при врахуванні випадковості необхідно щоби массові випадкові явища та події мали так звану властивість статичної стійкості. Це означає, що враховувані випадкові події підпорядковуються визначеним статичним закономірностям, вимоги котрих не обов’язкові при врахуванні невизначеності.

Умова статичної стійкості дозволяє використовувати в процесі прийняття рішень ефективні математичні методи теорії випадкових процесів, в тому числі один з її підрозділів – теорії марковських процесів.

Марковські процеси, зокрема марковські поля, дуже вдало використовуються для розв’язку різноманітних задач, які складно формалізувати. До таких задач відносяться і задачі соціології.

ОСНОВНА ЧАСТИНА

Теоретична частина

Розділ 1. Марковські процеси

Випадковий процес називається *марковським процесом* якщо для кожного моменту часу t імовірність будь-якого стану системи в майбутньому залежить тільки від її стану в теперішньому та не залежить від того, як система прийшла в цей стан.

Марковський процес зазвичай задають графом переходів із стану в стан. Вони мають два варіанти: графи з *дискретним* та *неперервним* часом.

$\xi(t), t \geq 0$ називається марковським процесом зі станами $x_1, \dots, x_s$ (число станів S є скінченим $S < \infty, S \geq 2$), якщо $\xi(t)$ приймає значення стану $x_1, \dots, x_s$ $\forall t \geq 0$ $\forall x_{i_1}, \dots, x_{i_n}$ $(\forall n = 3,4, \dots)$ $\forall t_1 < t_2 < \cdots < t_n$

$$P\{\xi(t_{n+1}) = x_{i_n+1} | \, \xi(t_n) = x_{i_n}, \dots, \xi(t_1) = x_{i_1}\} = P\{(\xi(t_n) = x_{i_n} \big| \, \xi(t_n) = x_{i_n}\}$$

Умовна функція розподілу

$$P\{\xi(t_{n+1}) = x_{i_{n+1}} | \, \xi(t_n) = x_{i_n}\} = F\{x_n, t_n; x_{n+1}, t_{n+1}\}$$

називається *марковською перехідною імовірністю*.

Марковські процеси можна поділити на класи в залежності від структури множини значень випадкового процесу *X* та множини часу *T*. Якщо множина *X* є дискретною, процес $\xi(t)$ називають *марковським ланцюгом*.

Якщо ж множина часу дискретна, то ми маємо справу з *марковським ланцюгом із дискретним часом*, якщо множина часу неперервна, то процес називається *ланцюгом Маркова з неперервним часом*.

Якщо обидві множини X та T є неперервними, то процес називається *неперервним марковським процесом*.

1.1 Марковський ланцюг із дискретним часом

Нехай в нас є випадковий процес $\xi(t)$, який приймає значення i = 1,2, … з кінцевої множини X.

$$\xi(t) = i, \xi(t') = j$$

Переходи із стану в стан відбуваються через деякі рівні проміжки часу, які називаються кроками.

$$|t' - t| = 1$$

Умовні вірогідності

$$P(\xi(t+1) = j | \xi(t) = i) = p_{ij}(t), \forall i, j \in X$$

утворюють матрицю імовірностей переходів ланцюга із стану в стан в конкретний момент часу t. Якщо імовірності переходів не залежать від часу, то ланцюг Маркова називається *однорідним* із матрицею імовірностей переходу за один крок

$$P = \begin{bmatrix} p_{11} & \cdots & p_{1s} \\ \vdots & \ddots & \vdots \\ p_{s1} & \cdots & p_{ss} \end{bmatrix}$$

$$p_{ij} > 0, \sum_{j=1}^{S} p_{ij} = 1 \ \forall i = 1, \dots, S$$

Таку матрицю називають *стохастичною*.

Для опису марковського ланцюга використовують граф ймовірностей переходу, вершини котрого описують можливі стани системи, стрілки від одної вершини до іншої показують можливі переходи між станами, а число над стрілкою позначає вірогідність переходу.

Наприклад, множина станів X має чотири стани $X = (1,2,3,4)$, матриця імовірностей переходу виглядає так:

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1 & 0 \\ 1/4 & 0 & 0 & 3/4 \end{bmatrix}$$

Тоді граф імовірностей переходу має такий вигляд:

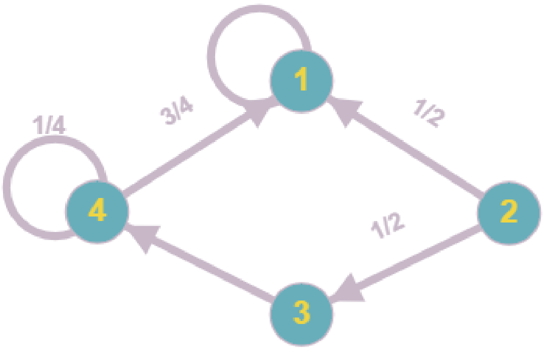

Рисунок 1

Ланцюги Маркова з дискретним часом визначаються *матрицею імовірностей переходу* за один крок з початковим розподілом

$$\Theta = \{\theta(1), \theta(2), \dots, \theta(n)\}, \text{де } \theta(i) = P\{\xi(0) = i\}.$$

Для розподілу імовірностей станів однорідного марковського ланцюга використовується рівність векторів:

$$\{p_1(k), p_2(k), \dots, p_n(k)\} = \{p_1(k-1), p_2(k-1), \dots, p_n(k-1)\} \cdot P = \Theta \cdot P^k$$

Для неоднорідного ланцюга використовується наступна формула для знаходження розподілу:

$$\{p_1(k), p_2(k), \dots, p_n(k)\} = \{p_1(k-1), p_2(k-1), \dots, p_n(k-1)\} \cdot P(k) \\ = \Theta \cdot P(1) \cdot P(2) \cdot \dots \cdot P(k)$$

1.2 Класифікація станів ланцюга Маркова з дискретним часом

Стан $i \epsilon X$ ланцюга Маркова називається *несуттєвим*, якщо $\exists\ m, j$:

$$p_{ij}{}^{(m)} > 0, \forall n\ p_{ij}{}^{(n)} = 0$$

Тобто існує такий стан j, в котре можна потрапити з додатною імовірністю, але з котрого неможливо повернутись в i. $p_{ij}{}^{(m)}$ – імовірність переходу ланцюга Маркова із стану *i* в стан *j* за *m* кроків.

Стан *j* називається *досяжним* із стану i $(i \rightarrow j)$, якщо $\exists n > 0$, що

$p_{ij}{}^{(n)} > 0$. Стани *j* та *i* називаються *сполучними* $(i \leftrightarrow j)$, якщо *j* досяжне з i та навпаки.

За визначенням відношення «↔» є симетричним $(i \leftrightarrow j\ \Rightarrow j \leftrightarrow i)$, та можна переконатись, що воно також і транзитивне $(i \leftrightarrow j,\ j \leftrightarrow k \Rightarrow i \leftrightarrow k)$.

Множину суттєвих станів можна розбити на скінченне число паралельних множин $X_1, X_2, \dots$, що складаються зі сполучних станів та характеризуються тим, що переходи між різними множинами неможливі.

Множини $X_1, X_2, \dots$ називаються *замкненими класами* або *нерозкладними класами* сполучних станів ланцюга Маркова. Марковський ланцюг, який утворює один нерозкладний клас, називається *нерозкладним*.

Стан замкненого класу j має період $d(j)$, якщо $d(j)$ – це найбільший спільний дільник чисел *n* таких, що:

$$p_{ij}{}^{(n)} > 0$$

Якщо $d(j) = 1$, $(d(X) = 1)$, то стан *j* класу *X* називається аперіодичним (ергодичним).

1.3 Класифікація станів ланцюга Маркова за асимптотичними властивостями перехідних імовірностей

Для ланцюга Маркова $\xi(n)$ визначимо

$$f_i(n) = P\{\xi(n) = i | \xi(0) = i, \xi(1) \neq i, \dots, \xi(n-1) \neq i\}$$

як імовірність першого повернення в стан i на кроці n, тоді $f_i = \sum_{n=1}^{\infty} f_i\,(n)$ – імовірність того, що система, яка вийшла зі стану *і* хоча б один раз повернеться в нього.

Стан $i \epsilon X$ називається *поворотним*, якщо $f_i = 1$, та *неповоротним*, якщо $f_i < 1$.

Всі стани кінцевого ергодичного класу поворотні. Неповоротні стани можливі тільки при нескінченному числі станів.

Якщо стан $i \epsilon X$ поворотний і $i \leftrightarrow j$, то стан $j \epsilon X$ також є поворотним.

Якщо стан i поворотний, тобто $f_i = 1$, то набір імовірностей $f_i(n)$ утворює розподіл імовірностей часу повернення.

Критерій поворотності станів. Стан $i \epsilon X$ є поворотним тоді і тільки тоді, коли

$$\sum_{n=1}^{\infty} p_{ij}(n) = \infty$$

Кожен поворотний стан можна віднести до одного з типів залежності від величини середнього значення часу повернення (скінченність або нескінченність). Величина $\mu_i = \sum_{n=1}^{\infty} n \cdot f_i(n)$ за визначенням математичного сподівання дорівнює середньому значенню кількості кроків, за котре ланцюг Маркова повернеться в стан *i*. Величина ${\mu_i}^{-1}$ характеризує інтенсивність повернення в стан *i*.

Поворотний стан i називається додатнім, якщо ${\mu_i}^{-1} > 0$, та нульовим, якщо ${\mu_i}^{-1} = 0$.

1.4 Марковський ланцюг із неперервним часом

Розглянемо марковський процес $\xi(t)$ з кінцевою множиною станів X, котрий не змінює свого стану в довільний момент часу. Такий процес називається ланцюгом Маркова з неперервним часом. Для цього марковського ланцюга виконуються умови

$P\{\xi(t) = j|\xi(s) = i, \xi(s') = i'\} = P\{\xi(t) = i|\xi(s) = j\} = p_{ij}(s,t),$
$\forall i, i', j \epsilon X, \forall s' < s < t \epsilon T$

Імовірність $p_{ij}(s,t)$ називається імовірністю переходу із стану i в стан j за проміжок часу *[s,t)*.

Ланцюги Маркова визначаються матрицей імовірностей переходу $P(s,t) = ||p_{ij}(s,t)||$ та початковим розподілом $\theta_i = P\{\xi(0) = i\}$.

Імовірності станів в будь-який момент часу t визначаються наступним чином:

$$P_j(t) = \sum_i \theta_i, p_{ij}(0,t)$$

Якщо імовірності переходів $p_{ij}(s,t)$ залежать тільки від різності моментів часу, тобто $p_{ij}(s,t) = p_{ij}(t-s) = p_{ij}(\tau)$, то ланцюг Маркова називається однорідним.

Для неоднорідних марковських ланцюгів імовірність переходу має вигляд $P(\tau) = ||p_{ij}(\tau)||$, а імовірності станів визначаються наступною формулою:

$$P_j(\tau) = \sum_i \theta_i, p_{ij}(\tau)$$

Перехідні вірогідності мають наступні властивості:

1) $p_{ij}(s,t) \geq 0$
2) $\sum_j p_{ij}(s,t) = 1$
3) Рівняння Чепмена – Колмагорова:

   $p_{ij}(s+t) = \sum_k p_{ik}(s) \cdot p_{kj}(t)$ – для однорідних ланцюгів Маркова

   $p_{ij}(s+t) = \sum_k p_{ik}(s,\tau) \cdot p_{kj}(\tau,t)$ – для неоднорідних ланцюгів Маркова, де $s < \tau < t$.
4) Умовні стохастичні неперервності:

$$\lim_{\tau \to 0} p_{ij}(\tau) = \begin{cases} 1, \text{якщо } i = j \\ 0, \text{якщо } i \neq j \end{cases}$$

Ця умова позначає, що з імовірністю 1 однорідний ланцюг Маркова не змінить свого стану за нескінченно малий проміжок часу $\tau \to 0$.

Розділ 2. Марковські випадкові поля

Нехай $\xi = \{\zeta_i | i \epsilon S\}, S = \{1, 2, \dots, N\}$ – багатовимірна випадкова величина, така, що кожна компонента ξ_j, що є одновимірною випадковою величиною, приймає значення x_j та визначена в своєму ймовірнісному просторі. Вважається, що $\forall j\ \ \xi_j$ є дискретними, визначені в одному ймовірнісному просторі та множина всіх значень зліченна. Введена таким чином величина ξ називається випадковим полем.

Конкретна реалізація $X = (x_1, \dots, x_N)$ багатовимірної випадкової величини ξ називається *конфігурацією*.

ξ – *випадкове поле* зі значеннями на множині $A = \{1,2, \dots, a\}, \forall i\ \xi_i \epsilon A$, якщо X – якась конкретна конфігурація ξ, то ζ – множина всіх можливих конфігурацій:

$$\zeta = \{X = (x_1, \dots, x_N) | x_i \epsilon A\ \forall i \epsilon S\}$$

Система сусідства – це множина $\partial = \{\partial_i | i \epsilon S\}$, де ∂_i – множина елементів з S, яке називається шаблоном сусідства для елемента i, таке що:

$$\begin{cases} i \notin \partial_i \\ i \in \partial_j \Leftrightarrow j \in \partial_i \end{cases}$$

Випадкове поле ξ називається *марковським випадковим полем* у відповідності до системи сусідства ∂ тоді і тільки тоді, коли $\forall i$:

$$\begin{cases} P(\xi = x) > 0\ \forall x \epsilon \zeta, \\ P(\xi_i = x_i | \xi_j = x_j, j \epsilon S \backslash \{i\}) = P((\xi_i = x_i | \xi_j = x_j, j \epsilon \partial_i) \end{cases}$$

Множина Y називається *клікою* тоді і тільки тоді, коли:

$$Y \subseteq y + \partial y, \forall y \in Y$$

Інакше кажучи, Y називається клікою тоді і тільки тоді, коли воно одноелементне (такі кліки ще називають тривіальними) або якщо кожен елемент Y є сусідом для кожного іншого елемента Y.

Нехай c – кліка, а x_c – обмеження конфігурації x на c, то є $x_c = (x_{i_1}, x_{i_2}, \dots, x_{i_{|c|}})$, де $i_j \epsilon c, j = 1 \dots |c|$. Нехай $C(\partial)$ – множина всіх клік для системи сусідства ∂, тоді потенційна функція $V_c(x_c)$ визначається як будь-яка функція V: $C(\partial) \to R$.

Дискретний розподіл називається розподілом Гіббса, якщо:

$$P(\xi = x) = \frac{1}{z} \exp(-\sum_{c \epsilon C(\partial)} V_c(x_c))$$

де Z – нормуюча константа, така що:

$$Z = \sum_{x \epsilon \zeta} \exp(-\sum_{c \epsilon C(\partial)} V_c(x_c))$$

Теорема Хаммерслея - Кліфорда: ξ – марковське випадкове поле тоді і тільки тоді, коли $P(\xi = x)$ – розподіл Гіббса.

Марковські процеси в соціології

Марковські процеси, зокрема марковські поля, дуже вдало використовуються для розв'язку різноманітних задач, які складно формалізувати. До таких задач відносяться і задачі соціології. На прикладі роздивимось математичну модель оцінювання суспільної думки.

Нехай населення країни можна описати деяким вектором параметрів $y = (y_1, y_2, \dots, y_m)$, кожний компонент котрого може бути як неперервним (наприклад, вік людини чи рівень її доходів), так і дискретним (наприклад, політичні погляди чи спеціальність). Можна вважати, що область значень неперервних величин ділиться на деяку скінченну кількість інтервалів та таких, що люди з характеристиками з одного з інтервалів ведуть себе однаково. Таким чином ми розділяємо населення на групи, які характеризуються фіксованим набором параметрів *y*. Наприклад, за віком ми можемо поділити людей на молодих, юних, дорослих, старих і т.д. Припустимо, що загальна кількість таких груп це *N*. Також припустимо, що суспільну думку теж можна поділити на деяку кількість можливих варіантів *M*. Тоді думку окремої людини можна представити як випадкову величину ζ, котра приймає значення від 1 до *M*. В такому випадку думка всього населення задається вектором $\xi = (\xi_1, \xi_2, \dots, \xi_N)$. Значення цього вектора приймають значення з множини можливих думок та поглядів *D*. Також ця думка може змінюватись із часом, тобто $\xi = \xi^t$ та залежати від значень параметрів вектора *y*. Існує вірогідність ${p_j}^{it}$, що представник суспільної групи *i* в деякий момент часу *t* займає позицію *j*.

$$ {p_j}^{it} = P\left({\xi_i}^t = j\right), \sum_{j=1}^{M} {p_j}^{it} = 1, {p_j}^{it} \geq 0 $$

Ми маємо суспільного діяча, що використовує деяку політику впливу для зміни суспільної думки. Цю політику можна охарактеризувати вектором параметрів $x^t = ({x_1}^t, {x_2}^t, \dots, {x_n}^t)$, який змінюється в часі. Ця політика впливу та ті заходи, які здійснюються для її підтримання, коштують грошей тож можна ввести функцію вартості $f(x^t)$.

Думка громадян змінюється в залежності від політики соціального впливу чи під впливом поглядів суспільства. Також є деякі групи суспільства, цінності яких є критичними для формування поглядів прошарку суспільства, позначимо такі критичні групи як $N(i)$, прошарок суспільства як i та задамо вектор поглядів $D_{N(i)}$.

Ми можемо визначити думки та погляди представника з прошарку i за допомогою умовного розподілу $H^i\left(D_i \middle| D_{N(i)}; x^t\right)$. Ці ж погляди можуть бути визначені за допомогою відповідних соціологічних опитувань, наприклад: якою буде ваша думка з приводу того чи іншого питання (від 1 до M), якщо відомі погляди інших груп суспільства ($D_{N(i)}$) та зовнішні параметри $x = (x_1, x_2, \dots, x_n)$? Такий розподіл визначає динаміку зміни поглядів відповідного представника деякого прошарку суспільства за допомогою відношення:

$$p_j{}^{it} = \sum_{D_{N(i)} \epsilon D} H^i(\xi_i{}^t = D_i \,|\, \xi_{N(i)}{}^{t-1} = D_{N(i)_i}; x^t) P(\xi_{N(i)}{}^{t-1} = D_{N(i)_i})$$

Погляди груп, де $N_i = \{\emptyset\}$, є представниками лише однієї групи та не входять в діапазон впливу інших груп, хоча самі можуть впливати на інші. Для таких груп умовний розподіл є безумовним.

Для остаточного визначення динаміки зміни станів системи потрібно зафіксувати початковий розподіл поглядів $P(\xi_i{}^t)$ для $t = 0$. Дане вище відношення дозволяє розрахувати $p_j{}^{it}$ для будь якого $t \geq 0$ якщо є дані для початкового моменту часу.

Наступним прикладом розглянемо математичну модель голосування на виборах. Ця модель дуже схожа на попередню за вхідними параметрами. Для розв'язку такої задачі було запропоновано загалом три моделі:

1. Модель Мартіна та Мея (1970 рік). Кожен виборець з самого початку має імовірність $\frac{1}{2}$ щоб проголосувати «так», «ні» і т.д. (знаменник збільшується від кількості варіантів проголосувати), зовнішнє поле відсутнє. Потім відносна імовірність того, що виборець проголосує «так», може стрімко зрости за рахунок фактору $e^{D/2}$ для кожного сусіда нашого виборця, який проголосував «так». Але відносна імовірність того, що виборець проголосує «так», може і впасти за рахунок іншого фактору $e^{-D/2}$, який дійсний для кожного сусіда, який проголосував «ні».
2. Модель Смуклера (1971 рік). Ця модель є динамічною. Уявимо, що ми маємо деяку взаємодію у часі *t*. В часі *t + 1* ми обираємо випадкового виборця. Він буде мати таку саму думку, як і більшість його сусідів (включаючи самого виборця), у момент часу t з імовірністю *1 – x*, де $0 \leq x \leq 1$.
3. Модель Кіндермана (1973 рік). Модель, в якій у момент часу t *+ 1* кожен виборець проголосує «так» з імовірністю
$$\frac{m^a}{m^a + m^b}$$
де m – це константа, a – число сусідів виборця, котрі проголосували «так» у час *t*, а b – число сусідів, що проголосували «ні».

Згодом, перша модель Мартіна та Мея була покращена та перевірена. У новій моделі всі сусіди могли впливати на дану імовірність виборця проголосувати «так», але нова модель вже не була Марковським полем. Також, всі три запропоновані моделі були перевірені на «марковість» та дали наступні результати: перша модель визначає розподіл Гіббса найближчого сусіда виборця, а тому є марковським полем; друга модель має часову умову відповідну до марковської та є процесом Маркова зі зворотним часом та з граничним розподілом кількості всіх «так», що є по суті біноміальним розподілом при $x \to \frac{1}{2}$ та розподілом типу «все або нічого» при $x \to 0$. Третя модель є прикладом синхронного процесу, тобто процесу, в якому кожен виборець голосує незалежно від думки інших виборців, які голосують в той самий час.

Ще один метод застосування Марковських полів в соціології був розроблений Холандом та Леінхардтом (1977 рік) та пізніше покращений Вассерманом (1977 рік). Вони використовували імовірнісний розподіл та стохастичні процеси для вивчення соціальних взаємодій, представляючи їх на графі G, де вектори $(a, b, \dots)$ представляли деяких осіб. Вершини цього графа представляли відношення між особами, наприклад, $a \to b$ може означати, що «a має вплив на b», «a захоплюється b», «a знає b» і т.д. Приклад такого графу:

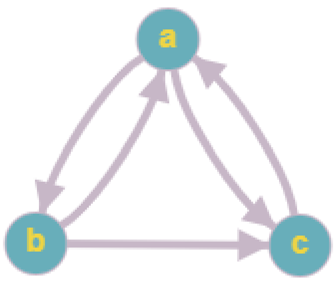

Рисунок 2

Кожен такий граф утворює суміжну матрицю. Наприклад, для даного графу:

$$X = \begin{array}{c} a \\ b \\ c \end{array} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Імовірнісний розподіл виглядає так:

$$p_l(x) = \left(\frac{l}{c}\right) \cdot e^{v_l \cdot n_m(x) + \sum_a v_a \cdot n_{a+}(x) + \sum_a \overline{v_a} \cdot n_{+a}(x)}$$

де $n_m(x)$ це кількість взаємних граней графу, $n_{a+}(x)$ – кількість граней, що виходять з вершини a, $n_{+a}(x)$ – кількість граней, що входять в вершину a, с – нормалізуюча константа та $v_l, v_a, \overline{v_a}$ – параметри. Такий розподіл може бути застосований для опису марковського випадкового поля (за Кіндерманом та Снеллом (1979 рік)).

Висновки

Основним завданням моєї роботи було дослідження застосування марковських ланцюгів та марковських полів в соціології. Найбільш вдалим їх використанням є розв’язок задач зі складною формалізацією, а більшість задач соціології є саме такими. За допомогою марковських полів можна створити доволі точну математичну модель взаємовідносин у суспільстві, суспільної думки та навіть модель проходження виборів та обрахувати вірогідності тієї чи іншої події з моделі.

Список використаної літератури

1. Волков И. К., Зуев С. М., Цветкова Г. М. Случайные процессы: учнебник для вузов. Издательство МГТУ им. Н.Э. Баумана, 1999. 26, 59 - 60, 163 - 169 с.
2. Кнопов П. С., Самосенок А. С. О некоторых прикладных задачах марковских случайных процессов с локальным взаимодействием. Статья, 2011. 19 - 23 с.
3. Kindermann R., Snell J. L. Markov Random Fields and Their Applications. American Mathematical Society, 1980. 24-33 с.
4. Kinderman R.P., Snell J.L. On the relation between Markov random fields and social networks. Journal of mathematical Sociologu, 1980,-vol.7, 1 - 13 с.
5. Hammersley J. M., Clifford P. Markov random fields in statistics. Unpublished paper, 1971. 21 - 22 с.
6. Назаров А. А., Терпугов А. Ф. Теория вероятностей и случайный процессов: учебное пособие. 2-е издание, исправленное. Издательство НТЛ, 2010. 41, 50 - 53, 109 – 111, 139 – 141, 155 - 157 с.
7. Дж. Кемени, Дж. Снелл Конечные цепи Маркова. Наука, М-1970г., 39 – 51 с.