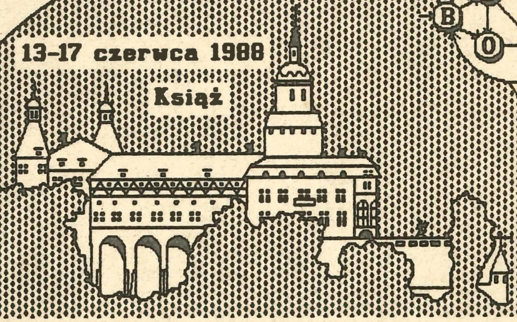
Redaktorzy:
A.Straszak
Z.Nahorski
J.Sikorski

13-17 czerwca 1988
Książ



1. Krajowa Konferencja Badań Operacyjnych i Systemowych

TOM 2

BOS'88

POLSKIE TOWARZYSTWO BADAŃ OPERACYJNYCH I SYSTEMOWYCH

INSTYTUT BADAŃ SYSTEMOWYCH POLSKIEJ AKADEMII NAUK

POLSKIE TOWARZYSTWO BADAŃ OPERACYJNYCH I SYSTEMOWYCH

TOM 2

WSPOMAGANIE PODEJMOWANIA DECYZJI
MODELE I SYSTEMY

I KRAJOWA KONFERENCJA

BADAŃ OPERACYJNYCH i SYSTEMOWYCH

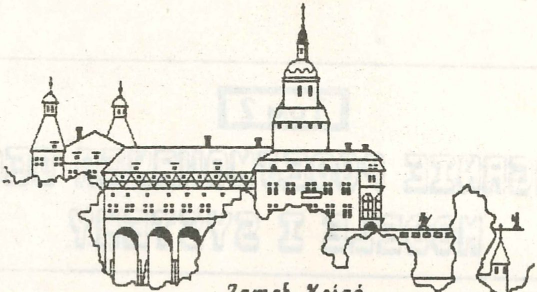
Książ, 13 - 17 czerwca 1988

BOS'88

INSTYTUT BADAŃ SYSTEMOWYCH POLSKIEJ AKADEMII NAUK

1989

WARSZAWA



Zamek Książ

I Krajowa Konferencja Badań Operacyjnych i Systemowych

Organizator konferencji

Polskie Towarzystwo Badań Operacyjnych i Systemowych
przy współpracy
Instytutu Badań Systemowych PAN

Komitet naukowy konferencji

Jerzy Hołubiec, Andrzej Kałuszko, Jerzy Kisielnicki, Henryk Kowalowski, Roman Kulikowski, Franciszek Marecki, Zbigniew Nahorski, Stanisław Piasecki, Jarosław Sikorski, Jan Stachowicz, Jan Stasierski, Andrzej Straszak, Maciej Sysło, Władysław Świtalski

Redaktorzy naukowi materiałów

Andrzej Straszak, Zbigniew Nahorski, Jarosław Sikorski

7. Systemy planowania i prognozowania

7.10

I Krajowa Konferencja
Badań Operacyjnych i Systemowych
Książ, 13 - 17 czerwca 1988r.

WYKORZYSTANIE ŁAŃCUCHÓW MARKOWA DO PROGNOZOWANIA PRZEMIAN STRUKTURY AGRARNEJ

Leszek Szydłowski
Instytut Badań Systemowych PAN
ul. Waryńskiego 17
71-310 Szczecin

W pracy wykorzystano metodę łańcuchów Markowa do sporządzania średniookresowych prognoz międzysektorowych przemian w zakresie władania użytkami rolnymi. Prognozy badanego procesu agrarnego zbudowano zarówno w oparciu o jednorodny jak i niejednorodny łańcuch Markowa. Szacowanie parametrów przeprowadzono w oparciu o makrodane wykorzystując metodę największej wiarygodności. Uzyskane wyniki, porównane z danymi rzeczywistymi dla dwóch lat realizacji prognozy, potwierdziły dużą trafność sporządzonych prognoz.

1.Wstęp

Na zjawisko międzysektorowych zmian własnościowego władania ziemią ma wpływ wiele czynników o charakterze prawnym, politycznym ekonomicznym i społeczno-kulturalnym. Wielość czynników oddziaływujących na dane zjawisko, a często także losowy ich charakter, skłaniają do przyjęcia założenia by traktować to zjawisko jako pewien układ stochastyczny a wielkości przez nie generowane jako zmienne losowe. Sugeruje to możliwość wykorzystania do opisu zmian procesów stochastycznych - łańcuchów Markowa. Proponowanym podejściem jest wykorzystanie macierzy prawdopodobieństw przejść jako transformant odzwierciedlających wpływ wszystkich czynników oddziaływujących w określonym przedziale czasu na badane zjawisko. Przyjmując założenie, że oddziaływanie czynników będzie kontynuowane w przyszłości z tą samą siłą, w oparciu o macierz prawdopodobieństw przejść sporządza się prognozę struktury rozważanego procesu agrarnego.

2.Metoda

Niech dana będzie skończona zbiorowość określonych jednostek (np.jednostkowych areałów ziemi) i pewna własność mierzalna W przysługująca poszczególnym jednostkom. Przez strukturę zbiorowości rozumiemy wektor y, którego składowe są frakcjami jednostek zaliczanych do wyodrębnionych przedziałów wartości cechy W (stanów). Zakładamy, że zarówno zbiorowość jak i jej struktura zależą od czsu t z pewnego przedziału <0, T>. Przyjmujemy, że obserwacje zbiorowości mogą się odbywać w dyskretnych momentach czasowych 0, 1, ..., T. Składowe $y_j(t)$ $(j=1, ..., r;\ t=0,...,T)$ wektora y(t) określone są następująco

$$y_j(t) = \frac{n_j(t)}{N(t)},\quad j=1,...,r;\ t=0,1,...,T,$$

gdzie $n_j(t)$ $(j=1, ...,r;\ t=0,1,...,T)$ jest liczbą jednostek znajdujących się w momencie t w stanie s_j, N(t) (t=0,1,...,T) jest liczbą wszystkich jednostek rozważanej zbiorowości w momencie t. Przez makrodane będziemy rozumieć zbiór wektorów y(t) (t=0, ...,T) rozważanej zbiorowości. y(t) są traktowane jako zmienne losowe.

Załóżmy, że proces przemian struktury badanej zbiorowości dobrze opisuje model jednorodnego łańcucha Markowa o macierzy prawdopodobieństw przejść P oszacowanej na podstawie danych z przedziału czasowego <0, T>. Przyjmując, że mechanizm przemian odzwierciedlony tą macierzą P nie ulegnie istotnym zmianom w okresie prognozowanym, prognozę rozważanej struktury wyznacza się w oparciu o zależność

$$y^{p}(t) = y(T)P^{t-T}, \qquad t>T$$

gdzie $y^{p}(t)$ jest wektorem prognozy struktury dla momentu t, a y(T) - wektorem opisującym empiryczną strukurę w momencie T, P jest oszacowaną macierzą prawdopodobieństw przejść odzwierciedlającą mechanizm przemian badanej zbiorowości.

W przypadku, gdy mechanizm przemian odwzorowuje ciąg macierzy prawdopodobieństw przejść P_s niejednorodnego łańcucha Markowa, który jest teoretycznym modelem opisującym zmiany struktury rozważanej zbiorowości, prognozę struktury wyznacza się z zależności

$$y^{p}(t) = y(T) \prod_{s=T+1}^{t} P_s, \qquad t>T$$

gdzie $y^{p}(t)$, y(T) są wektorami o znaczeniu jak wyżej, natomiast P_s jest macierzą ocen prawdopodobieństw przejść wyznaczoną w oparciu o zależności obrazujące zmiany prawdopodobieństw przejść w czasie.

Zakładamy, że zaobserwowane frakcje $y_j(t)$ $(j=1,...,r;$ $t=1,...,T)$ jednostek, które znalazły się w momencie t w stanie s_j wygenerowane zostały przez T zbiorów prób, każdy o N(t) niezależnych próbach. Przyjmujemy, że prawdopodobieństwo wpadania poszczególnej jednostki do poszczególnego stanu w obrębie danego zbioru prób jest stałe; zmienia się jednak wraz ze zmianą jednego zbioru prób na inny. Taki schemat probabilistyczny przedstawia rozkład wielomianowy. Tym samym w dowolnym t-tym zbiorze prób, w pojedyńczej próbie, jednostka może się znaleźć w każdym ze stanów s_1, s_2, ..., s_r z prawdopodobieństwem odpowiednio równym $q_1(t)$, $q_2(t)$, ..., $q_r(t)$, przy czym

$$\sum_{i=1}^{r} q_j(t) = 1 \quad \text{oraz} \quad q_j(t) = \sum_{i=1}^{r} y_i(t-1)p_{ij} \, .$$

Niech w momencie t zostanie wykonanych N(t) prób. Prawdopodobieństwo, że $n_1(t)$ jednostek znajdzie się w stanie s_1, $n_2(t)$ jednostek w stanie s_2, itd. i ostatecznie $n_r(t)$ jednostek w stanie s_r jest równe

$$\frac{N(t)! \prod_{j=1}^{r} q_j(t)^{n_j(t)}}{\prod_{i=1}^{r} n_i(t)!}, \quad \text{przy czym} \sum_{i=1}^{r} n_i(t) = N(t).$$

Łączna funkcja gęstości dla wyników niezależnych prób w czasie t (t=1, ...,T) jako funkcja parametrów prawdopodobieństw przejść przy określonym N(t) oraz zaobserwowanych frakcjach $y_i(t)$ jest funkcją wiarygodności.

Ponieważ prawdopodobieństwa przejść muszą spełniać określone warunki więc estymator p_{ij} oszacowany przy pomocy metody największej wiarygodności może być uzyskany drogą maksymalizacji funkcji wiarygodności przy warunkach

$$\sum_{j=1}^{r} p_{ij} = 1, \; i = 1, \ldots, r,$$

$$0 \leq p_{ij} \leq 1 \, , \; i,j = 1, \ldots, r.$$

Przy zastosowaniu warunków koniecznych Kuhn-Tucker'a znalezienie (ekstremum), oszacowań prawdopodobieństw p_{ij}, sprowadza się do rozwiązania następującego zadania programowania kwadratowego Lee i in. (1970):

zmaksymalizować

$$(X_*\Sigma_*^{-1}y_* - X_*\Sigma_*^{-1}X_*p_*)'p_* - \lambda' z_r \leq 0$$

przy ograniczeniach:

$$Rp_* \leq z_r ,$$

$$p_*, \lambda \geq 0,$$

gdzie p_* i y_* są odpowiednio $(r(r-1)\times 1)$ i $(T(r-1)\times 1)$ wymiarowymi wektorami, a X_* jest $(T(r-1)\times r(r-1))$ wymiarową macierzą postaci

$$p_* = \begin{bmatrix} p_1 \\ \dots \\ p_{r-1} \end{bmatrix}, \quad y_* = \begin{bmatrix} y_1 \\ \dots \\ y_{r-1} \end{bmatrix}, \quad X_* = \begin{bmatrix} X_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & X_{r-1} \end{bmatrix}, \text{ przy czym}$$

$$p_j = \begin{bmatrix} p_{1j} \\ \dots \\ p_{rj} \end{bmatrix}, \quad y_j = \begin{bmatrix} y_j(1) \\ \dots \\ y_j(T) \end{bmatrix}, \quad X_j = \begin{bmatrix} y_1(0) & \dots & y_r(0) \\ \dots & \dots & \dots \\ y_1(T-1) & \dots & y_r(T-1) \end{bmatrix}, \ j=1,\dots,r-1$$

$\Sigma_*^{-1} = [\Sigma^{ij}]$, $(i,j=1,\dots,r-1)$ jest $(T(r-1)\times T(r-1))$ wymiarową macierzą blokową o $(r-1)^2$ blokach będących $(T\times T)$ wymiarowymi macierzami diagonalnymi

$$\Sigma^{ij} = \begin{bmatrix} \frac{N(1)}{q_r(1)} & 0 & \dots & 0 \\ 0 & \frac{N(2)}{q_r(2)} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{N(T)}{q_r(T)} \end{bmatrix} \quad \text{dla } i \neq j$$

$$\Sigma^{jj}= \begin{bmatrix} \frac{N(1)}{q_j(1)}+\frac{N(1)}{q_r(1)} & 0 & \cdots & 0 \\ 0 & \frac{N(2)}{q_j(2)}+\frac{N(2)}{q_r(2)} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \frac{N(T)}{q_j(T)}+\frac{N(T)}{q_r(T)} \end{bmatrix},$$

z_r jest $(r\times1)$ wektorem którego składowe są jedynkami,

R jest $(r\times r(r-1))$ wymiarową macierzą złożoną z $(r-1)$ macierzy jednostkowych $(r\times r)$ wymiarowych,

λ jest wektorem mnożników Lagrange'a.

Macierz Σ_*^{-1} jest funkcją nieznanych wartości $q_j(t)$. Celem uzyskania estymatorów prawdopodobieństw przejść p_{ij} stosuje się następującą procedurę iteracyjną. Najpierw nieznany element $q_j(t)$ macierzy Σ_*^{-1} zastępowany jest przez zgodny (dostateczny i nieobciążony) estymator tj. frakcję $y_j(t)$, a następnie problem jest rozwiązywany celem znalezienia ocen parametrów p_{ij} oznaczanych jako $P^c(1) = [p^c_{ij}(1)]$. Dalej oblicza się

$$q^c_j(t) = \sum_{i=1}^{r} y_i(t-1)p^c_{ij}(1)$$

i wykorzystuje do oszacowania macierzy Σ_*^{-1}. Następnie znowu rozwiązuje się problem celem znalezienia $P^c(2) = [p^c_{ij}(2)]$, itd. Procedura ta jest powtarzana do momentu aż

$$\sum_{i=1}^{r}\sum_{j=1}^{r} |p^c_{ij}(n) - p^c_{ij}(n-1)| < \varepsilon,$$

gdzie ε jest zadaną wartością,

lub osiągnięta zostanie podana liczba kroków n.

Niech dany będzie ciąg (T+1) wektorów obserwacji strukury $y(t) = [y_1(t), y_2(t), \ldots, y_r(t)]$ (t=0, 1, ..., T).

Na postawie r+4 obserwacji struktury (tzn. wektorów y(t) (t=0, 1, ..., r+3) szacuje się macierz stałych prawdopodobieństw przejść $\tilde{P}_1$ wykorzystując procedurę estymacji omówioną wyżej. Macierz ta przyporządkowana jest następnie momentowi r+4, czyli uznaje się ją za macierz prawdopodobieństw przejść niejednorodnego łańcucha Markowa w okresie (r+3, r+4>. Następnie na podstawie wektorów struktury y(t) (t=1,..., r+4) szacuje się macierz $\tilde{P}_2$ przyporządkowując ją momentowi r+5, itd. Ostatecznie w oparciu o wektory y(t) (t=T-r-3, ..., T-1, T) szacuje się macierz $\tilde{P}_{T-r-2}$ którą przyporządkowuje się momentowi T+1.

W wyniku takiego postępowania uzyskuje się dla każdej pary (i, j) (i,j=1, 2, ..., r) T-r-2 wyrazowy ciąg ocen $\tilde{p}_{ij}(k)$ (k= 1, 2, ..., T-r-2) którego wyrazy odpowiadają kolejnym momentom t (t= r+4, r+5, ..., T+1). Następnie dokonuje się wyboru i oszacowania modeli tendencji rozwojowej oddzielnie dla poszczególnych prawdopodobieństw przejść. Oszacowane modele wykorzystuje się do wyznaczania ocen prawdopodobieństw przejść $p_{ij}(t)$ dla przyszłych momentów czsu.

3.Wyniki badań

Na podstawie danych za okres 1960-1983 r. ustalono 24 wektory y(t) (t=0, ..., 23) opisujące strukturę własnościową użytków rolnych w woj .szczecińskim. Składowe tych wektorów są frakcjami ziemi znajdującej się w jednym z następujących sektorów: państwowym, spółdzielczym i nieuspołecznionym.

Zakładając, że proces zmian struktury agrarnej opisuje model łańcucha Markowa dokonano oszacowania odpowiednich macierzy prawdopodobieństw przejść. Poniżej przedstawiono dwie macierze, jedną oszacowaną na podstawie 24 wektorów struktury (lata 1960 - 1983), drugą - na podstawie 11 wektorów (lata 1970-1984), dla zagadnienia użytkowania ziemi w sektorach Szydłowski (1987):

$$\hat{P}_{60-83} = \begin{array}{c} \\ (1) \\ (2) \\ (3) \end{array}\begin{array}{c} \begin{array}{ccc} (1) & (2) & (3) \end{array} \\ \begin{bmatrix} 0.9717 & 0.0283 & 0 \\ 0.1048 & 0.8049 & 0.0903 \\ 0.0346 & 0 & 0.9654 \end{bmatrix} \end{array}, \quad \hat{P}_{73-83} = \begin{array}{c} \begin{array}{ccc} (1) & (2) & (3) \end{array} \\ \begin{bmatrix} 0.9382 & 0.0386 & 0.0232 \\ 0 & 0.7811 & 0.2189 \\ 0.1060 & 0 & 0.8940 \end{bmatrix} \end{array}$$

$\chi^2 = 16,6147$, $\chi^2 = 4,1832$.

1, 2, 3 oznaczają odpowiednio sektor państwowy, spółdzielczy, nieuspołeczniony. Wartości statystyki χ^2 podane obok macierzy wskazują na dobre dopasowanie wektorów struktury, generowanych przez odpowiedni łańcuch Markowa (np. wektorów $y(t) = y(0)\hat{P}^t_{60-83}$) do odpowiednich empirycznych wektorów struktury (np. $y(t)$, $t=1, \ldots, 23$).

Prognozy struktury własnościowejużytków rolnych uzyskane na podstawie tych macierzy i wektora $y(23)$, opisującego strukturę w roku 1983, podane są w tab. 1.

Prognozę struktury użytkowania ziemi w sektorach wykonano także w oparciu o model niejednorodnego łańcucha Markowa. Ze względu na na to, że liczba stanów $r=3$ a $T=23$ możliwe było oszacowanie 18-tu macierzy $\tilde{P}_k$. Dla elementów $\tilde{p}_{ij}(k)$ $(i,j= 1,2,3;\ k=1,\ldots,18)$ przyjęto tendencję rozwojową postaci $\tilde{p}_{ij}(k)=ak^{-1}+b$. Następnie w oparciu o przyjętą tendencję rozwojową oszacowano macierze prawdopodobieństw przejść niejednorodnego łańcucha dla $t>23$. Wyniki prognoz podane są w tab. 1.

Celem oceny rzędu dokładności sporządzonych prognoz użytkowania gruntów osiągniętego w okresie empirycznej weryfikacji prognoz, tj. w latach 1984-85 obliczono błędy predykcji dla tych lat

$(U(t) = y(t) - y^P(t),\ t>T)$ oraz średnią arytmetyczną zaobserwowanych błędów predykcji U_s. Wielkości te podaje tab. 2.

Tabela 1

Prognoza zmian struktury użytkowania ziemi w sektorach woj. szczecińskiego do 1990 r

Sektor		LATA						
		1984	1985	1986	1987	1988	1989	1990
państwowy	a	0.5739	0.5782	0.5822	0.5860	0.5896	0.5930	0.5962
	b	0.5650	0.5665	0.5676	0.5683	0.5688	0.5691	0.5693
	c	0.5756	0.5860	0.5928	0.5976	0.6009	0.6032	0.6050
spółdzielczy	a	0.0824	0.0826	0.0829	0.0832	0.0835	0.0840	0.0844
	b	0.0911	0.0930	0.0945	0.0957	0.0967	0.0975	0.0981
	c	0.0954	0.0892	0.0863	0.0849	0.0844	0.0845	0.0847
nieuspołeczniony	a	0.3437	0.3392	0.3349	0.3308	0.3269	0.3230	0.3194
	b	0.3439	0.3405	0.3379	0.3360	0.3345	0.3334	0.3326
	c	0.3290	0.3248	0.3209	0.3175	0.3147	0.3123	0.3103

Prognoza w oparciu o model: j.ł M. a) $\hat{P}_{60-83}$, b) $\hat{P}_{73-83}$; c) n.ł.M.

Porównanie rzędów dokładności sporządzonych prognoz nie wykazało przewagi prognoz uzyskanych na podstawie modelu niejednorodnego łańcucha Markowa. Większa prostota jednorodnych łańcuchów Markowa, trafność i nieznaczna pracochłonność sprawia, że mogą one być cennym narzędziem pomocnym w planowaniu procesów agrarnych.

Należy zaznaczyć, że zaproponowane podejście może być wykorzystane do prognozowania innych zjawisk dla których dostępna jest informacja w postaci makrodanych.

Tabela 2

Wektory błędów predykcji U(24), U(25) i średniej arytmetycznej błędów predykcji U_s dla prognoz użytkowania ziemi w sektorach

		Sektor		
		państwowy	spółdzielczy	nieuspołeczniony
	a	-0.0110	0.0027	0.0083
U(24)	b	-0.0021	-0.0060	0.0081
	c	-0.0127	-0.0103	0.0230
	a	-0.0156	0.0019	0.0137
U(25)	b	-0.0039	-0.0085	0.0124
	c	-0.0234	-0.0047	0.0281
	a	-0.0133	0.0023	0.0105
U_s	b	-0.0030	-0.0073	0.0103
	c	-0.0181	-0.0075	0.0256

Obliczenia dla prognozy sporządzonej w oparciu o model: j.ł.M. o macierzy a) $\hat{P}_{60-83}$, b) $\hat{P}_{73-83}$; c) n.ł.M.

4. Literatura

1. Lee T.C., Judge G.G., Zellner A. (1970) Estimating the Parameters of the Markov Probability Model from Agregate Time Series Data. North-Holland Publishing Co.

2. Szydłowski L. (1987) Zastosowanie łańcuchów Markowa do badania i prognozowania procesów ekonomiczno-rolniczych zachodzących w regionie Pomorza Zachodniego. Praca doktorska, IBS PAN.

ZTI OIN PAN,zam.24/90,nakł.120egz.

Zarząd

Polskiego Towarzystwa Badań Operacyjnych i Systemowych

Prezes

prof.dr hab.inż. Andrzej Straszak
Instytut Badań Systemowych PAN

Wiceprezes

prof.dr hab.inż. Jan Stasierski
Wojskowa Akademia Techniczna

Wiceprezes

prof.dr hab.inż. Stanisław Piasecki
Instytut Badań Systemowych PAN

Sekretarz generalny

dr inż. Zbigniew Nahorski
Instytut Badań Systemowych PAN

Sekretarz

dr inż. Jarosław Sikorski
Instytut Badań Systemowych PAN

Skarbnik

dr inż. Andrzej Kałuszko
Instytut Badań Systemowych PAN

Członkowie

prof.dr hab. Jerzy Kisielnicki
Wydział Zarządzania UW

doc.dr hab.inż. Bohdan Korzan
Wojskowa Akademia Techniczna

doc.dr hab.inż. Jan Stachowicz
Zakład Nauk Zarządzania PAN

doc.dr hab.inż. Maciej Sysło
Instytut Informatyki UWr.

Komisja rewizyjna

PRZEWODNICZĄCY

dr Władysław Świtalski
Katedra Cybernetyki i Badań Operacyjnych UW

CZŁONKOWIE

dr inż. Janusz Kacprzyk
Instytut Badań Systemowych PAN

dr inż. Marek Malarski
Instytut Transportu PW

doc.dr hab.Henryk Sroka
Akademia Ekonomiczna w Katowicach

dr inż. Leon Słomiński
Instytut Badań Systemowych PAN