

# 토지이용 공간변화 예측의 통계학적 모형에 관한 연구

김의홍\*

## A Study on Statistical Modeling of Spatial Land-use Change Prediction

Eui-Hong Kim

### 요 약

토지이용 분류 체계상에서의 종류라는 개념은 토지이용 변화의 분류 체계상에 그대로 적용시킬 수가 있다. 본 연구에서는 선형 판별 함수를 원용하는 최우법(Maximum likelihood method)으로 산출되는 토지이용 분류의 공간적 결과와 Markov 전이 행렬 방법으로 산출되는 정량적 결과가 상호 보완하는 의미에서 합성 모형으로 통합되었다.

본 연구에서는 다변수 판별 함수의 계산법과 Markov 연쇄행렬 계산법에 관하여 토의되고 그 합성 모형을 대상 지역에 실제 적용하여 그 결과 '90년, '95년 토지이용도가 예측 작성되었다.

모형화의 문제 및 예측의 정확도 역시 더욱 토의 되어야 하며 추후 개선의 여지를 남긴다.

**ABSTRACT** : The concept of a class in the land-use classification system can be equally applied to a class in the land-use-change classification. The maximum likelihood method using linear discriminant function and Markov transition matrix method were integrated to a synthetic modeling effort in order to project spatial allocation of land-use-change and quantitative assignment of that prediction as a whole.

The algorithm of both the multivariate discriminant function and the Markov chain matrix were discussed and the test of synthetic model on the study area was resulted in the projection of '90 year as well as '95 year land -use classification.

The accuracy and the issue of modeling improvement were discussed eventually.

### 1. 서 론

본 연구의 목적은 최우법(maximum likelihood method)에 의한 토지이용 변화의 공간 할당 방법과 Markov 경향 예측 방법을 결합시킨 새로운 공간 예측의 모형을 수립하고 그 실증 과정에서 신뢰성과 문제점을 검토하는데 있다. 최우법과

Markov 방법을 결합시킨 이유는 공간 할당만으로 결과되는 한 방법과 계량적인 결과만을 산출하는 다른 한 방법의 결점을 상호 보완하기 위한 것이다.

일반적으로 분류(classification) 과정에서 정의되는 각 종류(class)는 토지이용의 범주에 적용되는 것과 똑같은 개념으로 토지이용변화의 범주에

\* 우성산업대학교 컴퓨터학과 (Department of Computer, Woo-Song University)

도 적용될수 있다. 그러므로 본 연구는 최우법에 의한 일정 시점의 토지이용 분류 방법을 2개 시점 간의 토지이용변화 분류에 그대로 확대 적용하고 그 결과를 Markov 변화 행렬과 합성시킨 일종의 통합 모형을 작성하는 방법론을 제시코자 하는 것이다.

모형 작성의 흐름도는 Fig. 1에서 보여주는 것과 같다.

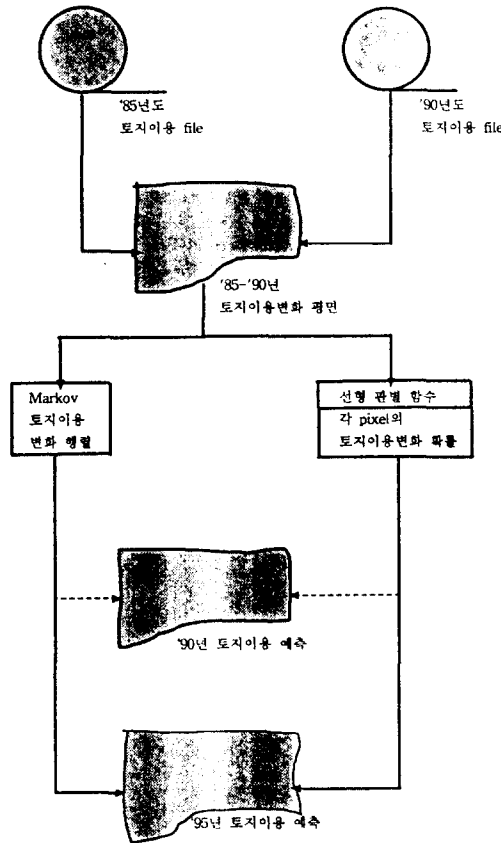


Fig. 1.1 토지이용 공간 변화 예측 모형 작성 흐름도

우선 기준 년도(85년)의 토지이용 현황과 참조 년도(90년)의 토지이용 현황 파일로부터 5년 간격(참조년도-기준년도)의 토지이용 변화를 검지 한다. 변화가 검지된 화소 자료를 이용하여 각 변화 중

류의 모든 경우를 대상으로 하는 적절한 판별 함수(discriminant function)를 작성한다. 그 함수를 '85년도 대상 지역의 각 화소에 적용, 각 화소의 토지이용변화의 사후 확률(posterior probability)을 산출한다. 각 토지이용으로의 변화 확률중 가장 높은 확률을 택하여 그 토지이용 종류로 지정 하는 것이 즉 최우법이라는 것은 주지의 사실이다.

한편 Markov 변화 행렬은 일정 시간 간격의 경과후 전체적인 각 토지이용 변화의 화소 수량을 산출 제공한다. 토지 전역을 구성하는 각 화소의 변화 확률들은 각 변화 종류별로 집합이 가능하고, 집합된 화소들은 그것들의 변화 발생의 사후 확률에 의거하여 고순위로 정렬(sort) 가능하다. 변화의 정수는 Markov 경향 행렬로부터 제공받은 수량만큼 고확률 순으로 선정된다. 각 화소의 정확한 공간 위치는 화면상의 주소(address)로 되돌아가 미래 특정 시점의 토지이용 예측도가 작성된다. 전 모형의 과정을 반복하면 일정 시간 간격으로 연속된 미래의 토지이용도 작성이 가능해진다.

## 2. Markov의 정량적 토지이용 예측

표 2.1은 '85-90년 제주도 일원의 토지이용변화 화소 수량의 행렬표이고 표 2.2는 각 행렬의 요소를 상대적인 빈도수로 변환시킨 것이다. 행렬의 전 행은 어느 행도 그 각 요소의 합계가 100%(확률 1)이 되는 stochastic vector이다. 즉  $Pr_{i1} = \sum_{j=1}^n Pr_{ij} = 1$ 이다. 토지이용변화 종류(class)별 화소 수 행렬이 Markov 행렬의 속성을 갖고 있는지의 여부 검증은 자유도  $(11-1)^2=100$ 으로 다음과 같이 시행되었다.

$$\chi^2 = 2 \sum_{i=1}^n \sum_{j=1}^n Tr_{ij} \ln \frac{Pr_{ij}}{Pr_j}$$

여기서  $Tr_{ij}$  는 변화 확률의 화소 수

$Pr_{ij}$  는 변화 확률 행렬 요소

표 2.1 토지이용 변화 화소 수량 행렬표('85-'90)

	urban	crop	orchar	broad	pastur	dense	sparse	barren	others	sea	Total
urban	314	26	16	17	128	12	8	97	27	1	646
crop	55	677	52	199	111	3	4	109	195	0	1405
orchar	15	45	136	114	36	104	46	23	43	167	729
broad	21	231	186	1102	52	144	77	33	164	0	2010
pastur	134	147	89	124	3019	55	137	268	467	0	4440
dense	2	9	70	99	17	790	670	3	6	0	1666
sparse	6	96	142	467	182	356	686	33	150	0	2118
barren	102	208	66	236	311	24	15	154	271	0	1387
others	80	392	97	322	298	27	9	136	367	0	1728
sea	0	1	108	1	0	0	0	0	0	4281	4391
Total	729	1832	962	2681	4154	1515	1652	856	1690	4449	20520

표 2.2 토지이용 변화 확률 행렬표('85-'90)

	urban	crop	orchar	broad	pastur	dense	sparse	barren	others	sea	Total
urban	48.6	4.0	2.5	2.6	19.8	1.9	1.2	15.0	4.2	0.2	100.0
crop	3.9	48.2	3.7	14.2	7.9	0.2	0.3	7.8	13.9	0.0	100.0
orchar	2.1	6.2	18.7	15.6	4.9	14.3	6.3	3.2	5.9	22.9	100.0
broad	1.0	11.5	9.3	54.8	2.6	7.2	3.8	1.6	8.2	0.0	100.0
pastur	3.0	3.3	2.0	2.8	68.0	1.2	3.1	6.0	10.5	0.0	100.0
dense	0.1	0.5	4.2	5.9	0.8	35.0	29.6	0.1	0.3	0.0	100.0
sparse	0.3	4.5	6.7	22.0	8.6	16.8	32.4	1.6	7.1	0.0	100.0
barren	7.4	15.0	4.8	17.0	22.4	1.7	1.1	11.1	19.5	0.0	100.0
others	4.6	22.7	5.6	18.6	17.2	1.6	0.5	7.9	21.2	0.0	100.0
sea	0.0	0.0	2.5	0.0	0.0	0.0	0.0	0.0	0.0	97.5	100.0

표 2.3 토지이용 변화 확률 행렬표('85-'95)

	urban	crop	orchar	broad	pastur	dense	sparse	barren	others	sea	Total
urban	25.7	8.2	3.6	7.9	27.8	2.9	2.6	10.9	8.9	0.8	100.0
crop	5.5	29.9	5.2	19.5	14.7	2.3	1.5	7.1	13.6	0.8	100.0
orchar	2.5	8.4	7.4	16.4	7.9	10.1	8.3	2.9	6.2	26.6	100.0
broad	2.3	14.9	8.4	36.3	6.9	8.7	6.2	3.3	9.3	2.1	100.0
pastur	4.6	7.8	3.3	8.0	50.7	2.6	3.9	6.4	11.7	0.5	100.0
dense	0.4	2.8	4.8	12.6	3.9	18.3	20.5	0.9	3.2	0.9	100.0
sparse	1.5	8.8	7.0	23.7	11.7	14.1	17.1	2.8	7.9	1.5	100.0
barren	6.8	16.7	5.4	18.9	24.5	3.6	2.9	6.8	12.8	1.1	100.0
others	5.5	20.2	5.7	20.4	20.6	3.6	2.5	6.5	13.1	1.3	100.0
sea	0.1	0.2	2.9	0.4	0.1	0.4	0.1	0.1	0.1	95.6	100.0

$P_j$  는 변화 확률 행렬 j열의 한계 확률

이 결과  $\chi^2=796972.00$ 으로 1%의 유의 수준으로도 귀무 가설 '어떤 변화 발생은 선행하는 상황에 종속되지않는다.'은 기각되었다. 그러므로 표 2.2 행렬은 Markov 연쇄의 속성을 갖는바, 즉 어떤 시점에서 어떤 사상의 상황은 확률적인 양태로서 선행하는 상황에 종속하여 변화한다. 이론적으로 Markov 확률 시스템에서는 현재의 상황으로부터 어떤 미래의 상황 예측도 가능하다.

Markov 연쇄의 변화 확률 행렬에서는 초기 시점으로부터 n 단계후의 확률의 vector는  $[Pr]n = [Pr]^n [P]n$  이 된다.

여기서  $[Pr]$ 은 초기 시점의 확률 vector,  $[P]$ 는 변화 확률 행렬이다. 본 연구에서  $[Pr]$ 은 표 2.2 의 각 행렬 vector,  $[P]$ 는 전 확률 행렬 vector로 대입하였다.

Markov 변화 행렬의 n 단계 후의 결과에는 초기 시점에 없는 요소도 나타나는 바, 본 연구에서는 그러한 요소는 변화 없음으로 가정하였다.

### 3. 선형 판별 함수(discriminant function)의 수립

본 연구의 판별 함수는 다음과 같이 다변수 다중류(class) 분류 함수  $S_{lmk} = C_{m0} + \sum_{j=1}^r C_{mj} X_{lkj}$  로 표현되고 여기서 class  $l=1,2,\dots,t$ , class  $m=1,2,\dots,g$ , 화소  $k=1,2,\dots,n$ , 변수  $j=1,2,\dots,r$  이다. 각 계수와 상수 그리고 값들은 아래와 같다.

a) 분류 함수의 각 변수별 계수와 상수

$$C_{ki} = (n-g) \sum_{j=1}^r x_{kj} a_{ij}$$

여기서 변수  $i=1,2,\dots,r$

class  $k=1,2,\dots,g$

화소  $j=1,2,\dots,n$

$$C_{k0} = -\frac{1}{2} \sum_{i=1}^r \bar{C}_{ki} x_{ki}$$

여기서  $k=1,2,\dots,g$

b) 각 쌍의 class 간의 Mahalanobis 거리의 2 제곱은

$$D_{ml}^2 = \sum_{i=1}^r (C_{mi} - C_{li}) (\bar{x}_{mi} - \bar{x}_{li})$$

여기서  $m, l=1,2,\dots,g$

c) 각 쌍의 class간의 상위성을 검증하는 F 값은

$$F_{ml} = (n-g-r+1)n_m n_l / r(n-g)(n_m+n_l) * D_{ml}^2$$

여기서 자유도 r과  $n-g-r+1$

class  $m, l=1,2,\dots,g$  이다.

변수로서는 '85년 LANDSAT MSS 4, 5, 6, 7 밴드의 화소 값 자료를 사용하였고, 마지막으로 다음과 같이 각 화소의 사후 확률(posterior probability)을 계산하였다.

여기서  $l=1,2,\dots,t$

$m=1,2,\dots,g$

$k=1,2,\dots,n$ 에 대하여

a) l class의 k 화소에 대한 m 번째 종류의 분류함수의 값

$$S_{lmk} = C_{m0} + \sum_{j=1}^r C_{mj} x_{lkj}$$

b) l class의 k 화소가 m class 로 분류될 사후 확률

$$P_{lmk} = p_m \exp(S_{lmk}) / \sum_{i=1}^m p_i \exp(S_{lik})$$

여기서  $P_m$ 은 m class의 사전 확률(prior probability) 이다.

#### 4. 통합(합성) 모형 실험 결과

판별 함수에 의한 분류 과정에서는 앞의 식에서 보여준 바와 같이 Mahalanobis 거리의 2제곱  $D^2$  을 사용한 F 값으로서 종류간의 상위성을 검증하면서, '85년 제주도 일원 전 대상 지역의 화소에 판별 함수를 적용하여 각 화소의 사후 확률을 계산하였다. 그것들을 각 변화 종류별로 집합시켜 변화 발생 확률의 고순위에 의거 정렬(sort)시켰다.

한편 각 변화 종류의 정수는  $[Pr1]=[Pr]*[P]1$  의 공식을 적용시켜 표 2.3 '85-'95 확률 행렬표를 작성하고 화소 수로 역산하여 산출하였다. 토지이용 변화 화소는 산출 제공된 수량에 의해 정렬되어 있는 목록의 고풍률 순에 따라 취해지고 각 화소는 자신이 간직하고있던 위치 주소(address)로써 화면의 원위치로 되돌아간다. 저확률의 화소는 변화 없음으로 간주되었다.

그러나 목록상의 화소수가 변화 행렬로부터의 수량보다 적은 경우, 보충분으로서 고 순위로 정렬시켜놓은 제2, 제3, ..., 제n의 사후 확률의 목록이 필요하게된다.

본 연구에서는 목록상의 화소 수가 부족한 경우, 부족분의 화소는 변화 없음으로 간주하였다.

'85년 토지이용 현황도는 도 4.1 과 같고 '90년 토지이용 현황도는 도 4.2 와 같고 '95년 토지이용 예측도는 도 4.3 과 같다.

본 논의에서는 생략되었으나, 모형에 의해 예측된 '90년 토지이용도 예측도와 '90년 토지이용 현황도를 대응 화소 별로 비교한 실증 결과 정확도가 높지않은 것은 대체로 다음과 같은 이유이다.

- 1) 본 실험에서는 전술한바와 같이 모형 작성의 방법론 제시에 의의를 두어 각 화소의 토지 이용변화 제1차 사후 확률만 적용하고 나머지 부족분에 대해서는 모두 변화없음으로 처리하였다.
- 2) 판별 함수의 변수로서 MSS의 각 밴드 값만

사용하였으나 그의 지형, 지문, 수문, 교통등의 GIS 자료를 보조하면 분류 정확도가 보다 향상될 것이다.

- 3) 사용된 MSS 자료는 그 해상도가 상대적으로 떨어지며 한국의 소규모 토지이용 현실과도 맞지않는다. 본 실험에서 사용된 '85년의 CCT는 3월 22일 주사된 가운데 한라산 정상 주위에 잔설이 있어 눈과 구름이 변화 종류의 traing 조로 혼입되어 그 통계적 오류가 분류과정에서 확산되었다.

#### 5. 결 론

본 연구에서는 토지이용변화 공간 예측의 한 방법론으로서, Markov 변화 행렬의 정량적 변화와 최우법(maximum likelihood)의 공간적 변화를 통합하여 일층 개량된 합성모형을 제시하였다.

본 연구의 실험 결과로서, 대체로 제주도 지역과 같이 토지이용의 변화가 자연 환경 요인에 의해 큰 영향을 받는 대상 지역에 대하여 이 방법이 유효하다. 그러나 본 실험에서 국한된 제주도 일원과 같이 제주공항 확충, 신시가지 개발등과 같은 의사 결정의 인문 환경 요인에 의해 큰 영향을 받는 대상 지역에 대하여는 판별 함수(discriminant function)의 변수로서 사회-경제적 요소의 추가가 바람직하다는 결론에 도달하였다.

본 예측 모형의 방법론은 토지이용(land-use) 변화 보다는 오히려 토지피복(land-cover) 변화에, 그리고 계절 분석(multi-date processing), 다중 분석(multi-file processing) 처리를 하는 농림, 수산, 수자원등의 GIS 타분야에 보다 효과적으로 적용될 수 있을 것으로 기대된다.

마지막으로 본 연구에서 최우법으로 적용한 다변수 다종류 선형 판별 함수 계산법(algorithm)은 일반화된 방법이긴 하나, 변화에 보다 적합한 계산법 고안의 개연성의 여지는 남겨 놓는다.

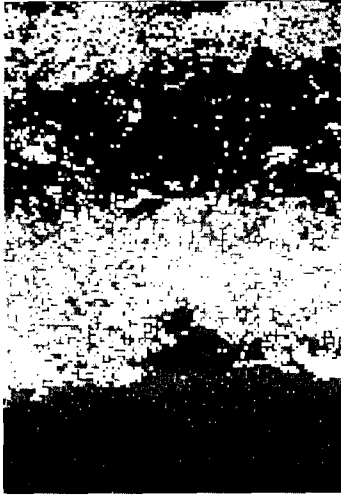


Fig. 4.1 '85년도 토지이용 현황도



Fig. 4.2 '90년도 토지이용 현황도



Fig. 4.3 '95년도 토지이용 예측도

## 참고 문헌

- 김의홍, 1995, 국토 정보 관리를 위한 리모트 센싱 기술 개발(3), 과학기술처, UCM20650
- 김의홍, 1994, 국토 정보 관리를 위한 리모트 센싱 기술 개발(2), 과학기술처, UCM20510
- 김의홍, 1993, 국토 정보 관리를 위한 리모트 센싱 기술 개발(1), 과학기술처, UCM20400
- 김의홍, 1984, 'A study on modeling of spatial land-use prediction', 生産研究36권8호, 東京大學 生産技術研究所
- LCWG, 1995, Proposed land-cover classification system, AARS
- Dale, V. H., R. V. O'Neill, M. Redlowski, and F. Southworth, 1993, 'Causes and Effects of Land-use change in Central rondonia, Brazil', Photogrammetric Engineering & Remote Sensing, Vol 59, No. 6
- Davis, F. W., and D. S. Simonett, 1991, GIS and Remote Sensing, Geographical Information System Principles and Applications
- Mater, P. M., 1987, Computer Processing of remotely sensed images, John Wiley & Sons
- Young, D., 1990, The X Window system programming and applications with Xt OSF/Motif Edition, Prentice-Hall