КЛАСИФИКАЦИЯ С ПОДМНОЖЕСТВЕНА СЕЛЕКЦИЯ НАПРЕД НА ПРИЗНАЦИТЕ ПРИ НОРМАЛНО РАЗПРЕДЕЛЕНИ ВЕКТОРИ

Димитър Ст. Тянев

Анотация: *Представена е процедура за подмножествена селекция напред на признаците за случай на два класа нормално разпределени вектори, изследвана чрез метода Монте-Карло в 54 варианта, осигуряваща оптимална вероятност за правилна класификация 0,90. Вариантите са характеризирани чрез сравняване на различни извадкови обеми, различни отношения на тези обеми и различни нива на темпа за достигане на Махаланобисовото разстояние 6,57. Показано е, че класификацията е зависима от съчетанието на тези параметри , които са непознати в практиката. Предлаганата двустепенна процедура за подмножествена селекция напред включва избор на споменатия темп преди прилагането на F-тестовете.*

CLASSIFICATION OF FORWARD SUBSET FEATURES WITH NORMALLY DISTRIBUTED VECTORS

Димитър Ст. Тянев

Abstract: *A procedure of forward subset features selection is presented for two classes of normally distributed vectors using Monte-Carlo algorithm in 54 versions, providing optimal propability of true classification equal to 0,90. These versions provide a comparison of different sampled volumes; different relations between the volumes and different levels of rate of securing the Mahalanobis distance equal to 6,57. It is shown that classification depends upon the combination of these parameters which are practically yet unknown. The two-stage procedure of forward subset selection covers the choice of the above stated rate before the F-tests run.*

Key words: Stepwise selection subset; Normally distributed vectors.

1. Въведение

Нека $\mathbf{x} = (x_1 x_2,..., x_k)^T$ е случаен вектор от k на брой измерени величини (признаци), който трябва да се класифицира към една от две възможни групи (класове) вектори $\Omega 1$ и $\Omega 2$, за които се знае, че са нормално разпреде-лени, с вектори на математическото очакване $\mu_i, i = 1,2$ и обща ковариационна матрица Σ. На практика параметрите на двете разпределения се оценяват със средните вектори $\mathbf{m}_i$ и претеглената ковариационна матрица $\mathbf{C}$, получени от статистическите извадки на двете групи, имащи обем n_i съответно. Тези реални параметри са елементи на линейната дискриминантна функция [1] от вида:

$$\text{(1)} \qquad d(\mathbf{x}) = (\mathbf{m}1 - \mathbf{m}2)^T.\mathbf{C}^{-1}.\mathbf{x} - \frac{1}{2}(\mathbf{m}1 + \mathbf{m}2)^T.\mathbf{C}^{-1}.(\mathbf{m}1 - \mathbf{m}2) .$$

Този вид на дискриминантната функция се получава при равновероятни групи (т.е. при априорна вероятност за събитието $\mathbf{x} \in \Omega 1$ и $\mathbf{x} \in \Omega 2$ равна на 1/2) Разпознаващото правило определя неизвестния вектор $\mathbf{X}$ като принадлежащ на първа група, ако $d(\mathbf{x}) \geq 0$ и като принадлежащ на втора група ако $d(\mathbf{x}) < 0$.

Тъй като оценките **m1**, **m2** и **C** са получени върху краен брой измервания, те оценяват параметрите $\mu 1$, $\mu 2$ и Σ с неизбежна грешка. Според Баесовия критерий, вероятностите за грешки при разпознаването се минимизират при нормални закони на разпределение, което позволява да бъдат оценени чрез една стандартна нормална комулативна функция на разпределение [2] както следва:

(2) $$\Pr(\Omega 2|\Omega 1) = \Pr(\Omega 1|\Omega 2) = \Phi(-\frac{\Delta}{2}) = \frac{1}{\sqrt{2\pi}} \int_{\Delta/2}^{\infty} e^{-\frac{\theta^2}{2}} d\theta ,$$

където

(3) $$\Delta^2 = (\mu 1 - \mu 2)^T . \mathbf{\Sigma}^{-1} . (\mu 1 - \mu 2)$$

е разстоянието на Махаланобис между двете групи.

Общата грешка от разпознаване се оценява с величината

(4) $$\varepsilon = \frac{1}{2} - \Phi(-\frac{\Delta}{2}) ,$$

от което следва извода, че колкото по-голямо е разстоянието между двете групи, толкова вероятността за грешка е по-малка.

В заключение трябва да се обърне внимание на изказаните съображения, тъй като линейната дискриминантна функция води при снижаване броя на наблюденията до снижаване и на вероятностите за правилна класификация. Този недостатък може да бъде силно изразен когато броят на измерваните величини k е по-голям в сравнение с обема на наблюдаваните извадки. Във всички случаи, когато се извършва класификация при неизвестни параметри на многомерните нормални разпределения, програмните средства получават статистическите им оценки. Един от резултатите на програмите за дискриминантен анализ е стойността на F-статистиката, която е една оценка на Махаланобисовото разстояние, тъй като статистическата му оценка е изместена [2]. Що се отнася до ковариационната матрица най-голяма нестабилност проявяват оценителите на корелациите, чийто брой е $\frac{k(k+1)}{2}$.

За да се облекчат споменатите пректически трудности класификацията може да бъде провеждана в едно от подпространствата на k-мерното признаково пространство. Може да се окаже, че линейната дискриминантна функция в подпространството има по-голяма вероятност за правилна класификация от тази в пълното k-мерно пространство. Въпреки, че са предложени много такива методи [2], вероятно най-често използувани са методите, основаващи се на постъпковия дискриминантен анализ, тъй като в големите програмни пакети по математическа статистика има подходящи алгоритми [2], [3], [4]. В случаите когато се изследват две групи тези методи се основават на серия от стандартни F-тестове в качеството им на оценители на Махаланобисовото разстояние. Две са основните версии на процедурите на постъпковия дискриминантен анализ, наричани селекция напред и елиминиране назад. Процедурата селекция напред започва с празно множество и с всяка следваща стъпка в линийната дискриминантна функция се въвежда ново измерване, докато по-нататъшните въвеждания на нови измервания престанат да водят до значимо (съществено) изменение на Махаланобисовото разстояние. Обратно, процедурата елиминиране назад започва с пълното множество като с всяка следваща стъпка се премахва едно измерване в линейната дискриминантна функция, докато не започне значително изменение в Махаланобисовото разстояние, оценявано върху оставащите измервания.

Когато общата статистика има малък обем N, N=n1+n2 в сравнение с размера на пространството k, началото на процедурата елиминиране назад може да бъде нестабилно (или невъзможно) поради горе посочените причини. Ето защо някаква версия на процедурата селекция напред е препоръчителна като работна. Необходимо е да се отбележи, че условието за прекратяване на работата на процедурата селекция напред не може да бъде формулирано строго за всички проблеми [5], където са показани сравненията между различни алгоритми за селекция напред при условие, че N е голямо в сравнение с k. Сравненията показват, че темпът на увеличаващото се Махаланобисово разстояние върху текущото съдържание на извадката е основният фактор за най-добрия алгоритъм. Целта на настоящият труд е да покаже, че подобни резултати са в сила и за случаите на малки по обем извадки (N<k) , както и да се предложи алтернативна процедура за селекция.

2. Методът селекция напред

Селекцията напред при класифициране в условията на две групи се осъществява най-добре чрез измененията в Махаланобисовото разстояние. Статистическата му оценка по аналогия с уравнение (3) има вида:

$$D^2_{(k)} = (\mathbf{m}1 - \mathbf{m}2)^T . \mathbf{C}^{-1} . (\mathbf{m}1 - \mathbf{m}2) \tag{5}$$

Нека сега $\mathbf{X}_{(p)} = (\mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \ldots, \mathbf{x}_{j_p})^T$ означава едно произволно **j**-то подмножество от **p** измервания (признака) $(p \le k), (\mathbf{X}_{(k)} \equiv \mathbf{X})$. Празното множество се означава $\mathbf{X}_{(0)}$. Съответстващите на $\mathbf{X}_{(p)}$ аналогични параметри с намалена размерност са $\mu 1_{(p)}$, $\mu 2_{(p)}$ и $\mathbf{\Sigma}_{(pp)}$, както и техните оценки $\mathbf{m}1_{(p)}$, $\mathbf{m}2_{(p)}$ и $\mathbf{C}_{(pp)}$.

Оценката на Махаланобисовото разстояние между формираните в подпространството извадки $\mathbf{X}_{(p)}$ се получава аналогично:

$$D^2_{(p)} = (\mathbf{m}1_{(p)} - \mathbf{m}2_{(p)})^T . \mathbf{C}^{-1}_{(pp)} . (\mathbf{m}1_{(p)} - \mathbf{m}2_{(p)}) \quad . \tag{6}$$

Този начин на изчисление осигурява истинно съответствие с Махаланобисовото разстояние $\Delta^2_{(p)}$.

Методът селекция напред се състои от две части:

1). Алгоритъм за селекция напред, който преобразува оригиналния вектор $\mathbf{x}_{(k)}$ в нов вектор $(x_{j_1} x_{j_2} \ldots x_{j_k})^T$. Разликата между двата вектора (в изходното k-мерно признаково пространство) е в подреждането на измерванията на параметрите. Същността на алгоритъма се изразява така: при известен вектор $\mathbf{X}_{(p-1)} = (\mathbf{x}_{j_1} \mathbf{x}_{j_2} \ldots \mathbf{x}_{j_{p-1}})^T$ за да се получи $\mathbf{X}_{(p)}$ трябва да се избере поредно измерване $\mathbf{x}_{j_p}$, което се добавя към вече избраните. Изборът се пада на онова измерване, което осигурява максимизация на $D^2_{(p)}$.

2). Условието за край в горния циклически алгоритъм, формиращ вектора $\mathbf{X}_{(p)}$ доставя най-добрата линейна дискриминантна функция в подпространството на избраните p на брой параметри.

За нашите цели бяха експериментирани седем нива на значимост като условие за край на F-включването F_α: α = 0,01; 0,05; 0,10; 0,15; 0,20; 0,25; 0,40; 0,70. Фактически се проверяваха нулевите хипотези $H0_{(p)}: \ \Delta^2_{(p+1)} = \Delta^2_{(p)}$ за всяко p**=**0, 1, 2, ...,(к-1), където $\Delta^2_{(0)} = 0$ за празното пространство $\mathbf{X}_{(0)}$.

Тестът на статистиката на двете групи се осъществява от величината [2]:

$$F_{(p)} = \frac{(n1+n2-p-2).n1.n2.\left(D^2_{(p+1)} - D^2_{(p)}\right)}{(n1+n2)(n1+n2-2)+n1.n2.D^2_{(p)}} \quad , \tag{7}$$

кедето $D^2_{(0)} = 0$.

Тази величина има F-разпределение относно нулевата хипотеза, която се отхвърля ако $F > F_{1-\alpha}$. Според това правило ще бъдат отхвърляни само силно корелираните измервания. Стойността F_α фактически избира онова $\mathbf{X}_{(p)}$, чието F-включване е максимално, но само ако съответното му $F_{(p)} \geq F_{1-\alpha}(1, n1+n2-p-2)$, при степени на свобода 1 и (n1+n2-p-2). Всяка друга стойност на α трябва да бъде разбирана като недействителна.

3. Вероятности за правилна класификация при селекция напред

При положение, че чрез горния алгоритъм е формиран един набор от най-добрите p на брой измервани величини, които формират векторите $\mathbf{X}_{(p)}$ в отсечените по този начин извадки, то аналогично на уравнение (1) линейната дискриминантна функция сега има вида:

$$d_{(p)}(\mathbf{X}_{(p)}) = (\mathbf{m}1_{(p)} - \mathbf{m}2_{(p)})^T.\mathbf{C}^{-1}_{(pp)}.\mathbf{X}_{(p)} - \frac{1}{2}(\mathbf{m}1_{(p)} + \mathbf{m}2_{(p)})^T.\mathbf{C}^{-1}_{(pp)}.(\mathbf{m}1_{(p)} - \mathbf{m}2_{(p)}) \tag{8}$$

и $\mathbf{X}_{(p)}$ се класифицира според същото правило.

Тъй като ковариационната матрица (като дясна нормална матрица) е положително определена, то $\Phi(\Delta_{(k)}/2) < \Phi(\Delta_{(k)}/2)$, т.е. наборът от измервания включен в $\mathbf{X}_{(p)}$ е възможно най-добрият избор.

За провеждане на класификацията бяха използвани още условните вероятности за правилна класификация. При известни $\mathbf{m}1_{(p)}$, $\mathbf{m}2_{(p)}$ и $\mathbf{C}_{(pp)}$, получавани чрез алгоритъма за селекция напред, условната вероятност за правилна класификация за групите Ωi е следната:

$$\Pr Ci_{(p)} = \Phi\left[(-1)^{i+1}.\mathbf{d}_{(p)}\left(\mu i_{(p)}\right)/\sigma_{(p)}\right], \; i = 1,2 \tag{9}$$

където

$$\sigma^2_{(p)} = (\mathbf{m}1_{(p)} - \mathbf{m}2_{(p)})^T.\mathbf{C}^{-1}_{(pp)}.\mathbf{\Sigma}_{(pp)}.\mathbf{C}^{-1}_{(pp)}.(\mathbf{m}1_{(p)} - \mathbf{m}2_{(p)}) \quad , \tag{10}$$

а векторите $\boldsymbol{\mu}\mathbf{i}$ и матрицата Σ са параметрите на генераторите на многомерните извадки.

Общата вероятност за двете групи се оценява чрез усредняване

$$\Pr \mathbf{C} = \left(\Pr \mathbf{C}1 + \Pr \mathbf{C}2\right)/2 \tag{11}$$

Друга уместна мярка представляват безусловните вероятности за правилна класификация, които за извадките формирани чрез алгоритъма за селекция напред се определят така:

$$\Pr \mathbf{U}i_{(p)} = \mathrm{E}\left\{\Pr \mathbf{C}i_{(p)}\right\} \quad , \quad i = 1,2 \tag{12}$$

където вероятността се отнася към контролни извадки с обеми n1 и n2 съответно. Общата безусловна вероятност се дефинира отново чрез усредняване:

$$\text{(13)}\qquad \Pr \mathbf{U}_{(p)} = \left(\Pr \mathbf{U}1_{(p)} + \Pr \mathbf{U}2_{(p)}\right)/2 \quad .$$

Да предположим, че дадено условие S за край на селекцията избира $\mathbf{X}_{(p)}$ като най-добро подмножество. Свързаните групови и обща условни вероятности се дефинира така:

$$\Pr \mathbf{C}i(S) \equiv \Pr \mathbf{C}i_{(p)} \;;\quad i = 1,2\;;\qquad \Pr \mathbf{C}(S) \equiv \Pr \mathbf{C}_{(p)} \quad .$$

При всяка от горните вероятности включени в условието S би се изменял както съставът така и обемът на селектираното подмножество от параметри.

4. Организация на експеримента

Според метода Монте-Карло извадките бяха генерирани като многомерни нормално разпределени с параметри $\boldsymbol{\mu}1 = 0$, $\boldsymbol{\mu}2 \neq 0$ (с положителни елементи) и $\boldsymbol{\Sigma}_{(kk)}$. При това елементите на вектора $\boldsymbol{\mu}2$ бяха определяни чрез следната геометрична параметризация [5]:

$$\text{(14)}\qquad \boldsymbol{\mu}2_p = \left[6{,}57(1-A).A^{p-1}/(1-A^k)\right]^{1/2},\ p = \overline{1,k}$$

където A се избира в отворения единичен интервал. По този начин елементите на вектора $\boldsymbol{\mu}2$ се подреждат в низходящ ред.

В резултат на това прилагането на алгоритъма за селекция напред ще може да възпроизвежда в естествен ред разстоянията $\Delta^2_{(p)} = \sum_{i=1}^{p} (\mu 2_i)^2,\ p = \overline{1,k}$. Тази формулировка съдържа в себе си пълното множество от параметри $\mathbf{X}_{(k)}$, което има Махаланобисово разстояние 6,57. Тази стойност на разстоянието е заложена в генерираните извадки за да се осигури вероятност за правилна класификация на ниво 0,90. Тук направеното проучване се основава на 54 извадкови ситуации, които са определени от всички възможни комбинации на споменатите параметри, обединени в таблица 1.

Таблица 1.

Описание на параметъра	Назначавани стойности
Темп на увеличаване на разстоянието $\Delta^2_{(p)}$	A= 0,275; 0,800; 0,975
Размерност на пространството	k= 10; 20; 30
Общ обем на извадките (n1+n2)	N= k+2; 2(k+2)
Отношение на обема на извадките (n1/n2)	R= 1; 2; 5

Избраните стойности на параметъра A осигуряват на средния вектор $\boldsymbol{\mu}2$, според зависимостта (14), различен темп на достигане на условието за край в алгоритъма за селекция. Така според назначаваните стойности този темп се определя като "бърз", "среден" и "бавен".

Получените резултати се основават на 100 варианта от всяка извадкова ситуация. Използувани са генераторите на многомерно нормално разпределение, изследвани в [6]. Оценките за вероятностите PrU(S), броят на селектираните параметри k(S) и тяхните средни отклонения (со) са усреднени върху общия брой варианти за всяка извадкова ситуация. Резултатите върху извадките в селектираното подмножество (означени с Sp) за трите вида параметри A, k, и R са дадени в таблици 2, 3, и 4 съответно, като за сравнение освен тях е даден и резултатът за пълното множество параметри (означен с Sk).

Таблица 2.

N	S	A = бързо PU(S)(co)	A = бързо k(S) (co)	A = средно PU(S)(co)	A = средно k(S) (co)	A = бавно PU(S)(co)	A = бавно k(S) (co)
k+2	Sp	0,851 (,047)	1,9 (1,8)	0,763 (,058)	6,2 (3,7)	0,739 (,043)	8,6 (3,8)
	F,70	0,653 (,078)	17,9 (2,1)	0,634 (,086)	18,8 (2,1)	0,634 (,082)	17,8 (2,1)
	F,40	0,695 (,081)	12,3 (3,5)	0,675 (,077)	12,5 (3,4)	0,681 (,076)	12,8 (3,1)
	F,25	0,734 (,075)	8,1 (3,2)	0,678 (,071)	9,0 (3,0)	0,694 (,058)	9,1 (3,0)
	F,15	0,763 (,076)	5,3 (2,5)	0,704 (,065)	6,1 (2,3)	0,697 (,049)	6,2 (2,4)
	F,10	0,776 (,079)	4,1 (2,0)	0,712 (,058)	4,6 (1,9)	0,687 (,047)	4,8 (1,9)
	F,05	0,791 (,083)	2,5 (1,2)	0,709 (,057)	3,0 (1,4)	0,664 (,048)	3,1 (1,5)
	F,01	0,787 (,105)	1,2 (0,7)	0,693 (,067)	1,4 (0,8)	0,596 (,061)	1,2 (0,9)
	Sk	0,612 (,091)	20,0 (---)	0,611 (,082)	20,0 (---)	0,609 (,085)	20,0 (---)
2(k+2)	Sp	0,873 (,015)	2,0 (1,2)	0,817 (,030)	9,9 (4,4)	0,807 (,031)	14,2 (3,7)
	F,70	0,793 (,037)	15,2 (2,0)	0,791 (,041)	15,9 (1,8)	0,791 (,039)	16,3 (1,7)
	F,40	0,811 (,037)	10,1 (2,3)	0,792 (,039)	11,7 (2,1)	0,786 (,038)	12,4 (2,1)
	F,25	0,821 (,041)	7,2 (2,1)	0,789 (,038)	9,1 (2,1)	0,776 (,035)	10,1 (2,0)
	F,15	0,827 (,040)	5,2 (1,9)	0,786 (,038)	7,1 (1,8)	0,763 (,033)	8,0 (2,0)
	F,10	0,841 (,038)	4,0 (1,5)	0,781 (,039)	5,8 (1,7)	0,748 (,034)	6,6 (1,8)
	F,05	0,849 (,037)	2,8 (1,2)	0,771 (,037)	4,3 (1,4)	0,727 (,033)	5,1 (1,5)
	F,01	0,855 (,036)	1,7 (0,7)	0,730 (,058)	2,4 (1,1)	0,670 (,041)	2,6 (1,1)
	Sk	0,791 (,039)	20,0 (---)	0,792 (,040)	20,0 (---)	0,793 (,040)	20,0 (---)

Таблица 3.

N	S	P = 10 PU(S)(co)	P = 10 k(S) (co)	P = 20 PU(S)(co)	P = 20 k(S) (co)	P = 30 PU(S)(co)	P = 30 k(S) (co)
k+2	Sp	0,781 (,055)	4,0 (2,2)	0,778 (,044)	5,6 (3,4)	0,781 (,033)	6,9 (3,8)
	F,70	0,663 (,102)	8,4 (1,5)	0,629 (,082)	17,7 (2,2)	0,613 (,072)	27,3 (2,6)
	F,40	0,698 (,083)	5,9 (2,0)	0,683 (,073)	12,6 (3,4)	0,675 (,066)	19,1 (4,7)
	F,25	0,722 (,075)	4,3 (1,8)	0,709 (,066)	8,6 (3,1)	0,701 (,059)	13,2 (4,4)
	F,15	0,723 (,071)	3,0 (1,5)	0,724 (,063)	5,8 (2,4)	0,730 (,052)	8,8 (3,4)
	F,10	0,716 (,073)	2,3 (1,2)	0,730 (,058	4,5 (1,9)	0,733 (,051)	6,6 (2,6)
	F,05	0,691 (,087)	1,6 (0,9)	0,721 (,059)	3,0 (1,4)	0,731 (,046)	4,1 (1,7)
	F,01	0,627 (,109)	0,7 (0,6)	0,710 (,081)	1,3 (0,8)	0,713 (,051)	1,8 (1,0)
	Sk	0,632 (,103)	10,0 (---)	0,608 (,080)	20,0 (---)	0,587 (,071)	30,0 (---)
2(k+2)	Sp	0,833 (,031)	5,4 (1,8)	0,834 (,023)	9,3 (3,0)	0,838 (,023)	11,5 (4,5)
	F,70	0,798 (,047)	8,0 (1,2)	0,797 (,038)	15,8 (1,8)	0,791 (,035)	23,5 (2,3)
	F,40	0,798 (,047)	5,9 (1,5)	0,801 (,039)	11,5 (2,1)	0,795 (,034)	16,7 (2,8)
	F,25	0,795 (,041)	4,7 (1,4)	0,799 (,037)	9,0 (2,0)	0,795 (,033)	12,6 (2,8)
	F,15	0,797 (,044)	3,7 (1,3)	0,796 (,035)	6,8 (1,8)	0,796 (,038)	9,6 (2,4)
	F,10	0,787 (,042)	3,1 (1,2)	0,791 (,034)	5,6 (1,7)	0,793 (,030)	7,9 (2,1)
	F,05	0,773 (,043)	2,4 (1,0)	0,779 (,031)	4,2 (1,4)	0,784 (,031)	5,7 (1,7)
	F,01	0,726 (,066)	1,4 (0,7)	0,761 (,033)	2,3 (1,1)	0,771 (,032)	3,0 (1,1)
	Sk	0,795 (,042)	10,0 (---)	0,801 (,038)	20,0 (---)	0,782 (,032)	30,0 (---)

Таблица 4.

N	S	R = 1 PU(S)(co)	R = 1 k(S) (co)	R = 2 PU(S)(co)	R = 2 k(S) (co)	R = 5 PU(S)(co)	R = 5 k(S) (co)
k+2	Sp	0,792 (,037)	5,5 (2,8)	0,783 (,044)	5,5 (3,1)	0,769 (,049)	5,6 (3,4)
	F,70	0,643 (,078)	17,9 (2,1)	0,640 (,085)	17,9 (2,1)	0,620 (,087)	17,8 (2,1)
	F,40	0,698 (,073)	12,3 (3,1)	0,691 (,071)	12,7 (3,5)	0,671 (,078)	13,2 (3,1)
	F,25	0,701 (,075)	8,5 (3,2)	0,717 (,063)	8,9 (3,1)	0,694 (,071)	8,6 (3,0)
	F,15	0,723 (,066)	6,3 (2,5)	0,730 (,058)	6,0 (2,4)	0,711 (,068)	5,7 (2,2)
	F,10	0,739 (,059)	4,1 (2,0)	0,729 (,058)	4,6 (1,9)	0,712 (,069)	4,3 (2,0)
	F,05	0,742 (,057)	3,5 (1,4)	0,722 (,065)	3,1 (1,4)	0,693 (,073)	2,9 (1,4)
	F,01	0,737 (,057)	1,7 (0,9)	0,692 (,079)	1,4 (0,8)	0,651 (,083)	1,1 (0,8)
	Sk	0,622 (,087)	20,0 (---)	0,617 (,081)	20,0 (---)	0,597 (,091)	20,0 (---)
2(k+2)	Sp	0,838 (,025)	8,8 (3,1)	0,838 (,027)	8,7 (2,9)	0,832 (,031)	8,7 (3,3)
	F,70	0,793 (,037)	15,9 (1,8)	0,798 (,037)	15,8 (1,8)	0,801 (,039)	15,7 (1,8)
	F,40	0,811 (,034)	11,6 (2,2)	0,802 (,037)	11,4 (2,1)	0,779 (,043)	12,1 (2,3)
	F,25	0,808 (,036)	9,0 (2,1)	0,804 (,036)	8,9 (2,0)	0,778 (,044)	8,6 (2,1)
	F,15	0,807 (,033)	7,0 (1,9)	0,801 (,034)	6,8 (1,9)	0,776 (,046)	6,5 (2,0)
	F,10	0,803 (,032)	5,8 (1,7)	0,798 (,033)	5,7 (1,6)	0,772 (,043)	5,1 (1,7)
	F,05	0,796 (,031)	4,5 (1,2)	0,789 (,031)	4,2 (1,3)	0,763 (,043)	3,7 (1,3)
	F,01	0,763 (,036)	2,4 (0,9)	0,769 (,040)	2,4 (1,1)	0,726 (,059)	1,9 (0,7)
	Sk	0,805 (,039)	20,0 (---)	0,798 (,038)	20,0 (---)	0,778 (,044)	20,0 (---)

Тези резултати водят до следните заключения:

а) във всички случаи селекцията напред чрез използуване на някакво F_α осигурява значително намаляване на обема на подмножеството измервани параметри. При това очакваната класификация е с около 30% по-добра от тази, която се постига с пълния набор параметри $\mathbf{X}_{(k)}$. Като цяло класификацията се подобрява с нарастване на N;

б) всички правила се изявяват най-добре когато темпът А се оценява като "бърз";

в) най-добрите правила при А≡"бързо" се оказват най-лошите при А≡"бавно";

г) принципните ефекти от нееднаквия обем на двете извадки (R≠1) се изразяваха в снижаване на процента на правилна класификация при всички правила;

д) препоръчва се условието за край на алгоритъма за селекция F_α да работи със следните нива на значимост:

		α	$\leq$	0,05	за бърз темп;
0,10	$\leq$	α	$<$	0,25	за среден темп;
0,25	$\leq$	α			за бавен темп.

Литература

[1]. Tou J.T., Gonzalez R.C., - *Pattern recognition principles*, Addison-Wesley pub. company, London, 1974.

[2]. Афифи А., Эйзен С., - *Статистический анализ. Подход с использованием ЭВМ.*, Москва, Издат. "Мир", 1982.

[3]. Jennrich R.I., - *Stepwise discriminant analysis. Statistical methods for digital computers*., (Edited by K. Enslein, A. Ralston and H.S. Wilf), Wiley, New York, 1977.

[4]. *Справочник по прикладной статистике*, том 1 и том 2, Под ред. Э. Ллойда, У.Ледермана, Ю.Н.Тюрина, - Москва, Издат. "Финансы и статистика",1989.

[5]. Constanza M.C., Afifi A.A., - *Comparison of stopping in forward stepwise discriminant analysis*., J. Amer. Stat. Assoc., 47, p. 777, 1979.

[6]. Тянев Д.С., - *Алгоритми за разпознаване на образи и използуването им в техническата диагностика*, Дисертация, Варна, 1990.