

Лабораторная работа

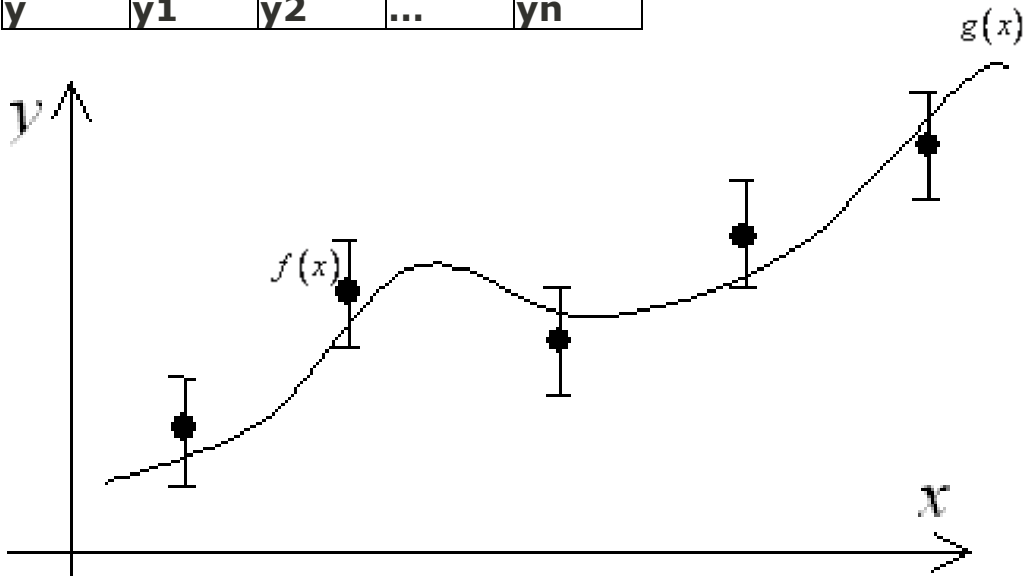
Тема: “Аппроксимация. Метод наименьших квадратов (МНК).”

Цель работы: научиться строить аппроксимирующие функции методом наименьших квадратов, оценивать их близость к исходной функции, использовать линии тренда.

Аппроксимацией (приближением) функции $f(x)$ называется нахождение такой функции $g(x)$ (аппроксимирующей функции), которая была бы близка заданной. Критерии близости функций могут быть различные.

f(x)

x	x1	x2	...	xn
y	y1	y2	...	yn



Подбор эмпирических формул.

Построение эмпирической формулы состоит из двух этапов: подбора вида этой формулы, содержащей неизвестные параметры и определение наилучших в некотором смысле этих параметров.

Вид формулы иногда известен из физических соображений или выбираются из геометрических соображений: экспериментальные точки наносятся на график и примерно угадывается общий вид зависимости путем сравнения полученной кривой с графиками известных функций

Для практики важен случай аппроксимации функции многочленами, т.е.

$$F(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_m x^m$$

Метод наименьших квадратов

Пусть для исходных данных $x_i, f_i, i=1, \dots, N$ (нумерацию лучше начинать с единицы), выбран вид эмпирической зависимости:

$$y = \varphi(x, a_0, a_1, \dots, a_m)$$

с неизвестными коэффициентами

a_0, a_1, \dots, a_m . Запишем сумму квадратов отклонений между вычисленными по эмпирической формуле и заданными опытными

данными:

$$S(a_0, a_1, \dots, a_m) = \sum_{i=1}^N (\varphi(x_i, a_0, a_1, \dots, a_m) - f_i)^2$$

Параметры a_0, a_1, \dots, a_m будем находить из условия минимума функции $S(a_0, a_1, \dots, a_m)$. В этом состоит метод наименьших квадратов (МНК).

Известно, что в точке минимума все частные производные от S по a_0, a_1, \dots, a_m равны нулю:

$$\frac{\partial S}{\partial a_0} = 0, \quad \frac{\partial S}{\partial a_1} = 0, \quad \dots, \quad \frac{\partial S}{\partial a_m} = 0 \tag{1}$$

Рассмотрим применение МНК для частного случая, широко используемого на практике. В качестве эмпирической функции рассмотрим

полином $\varphi(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_m x^m$

Формула для определения суммы квадратов отклонений примет вид:

$$S(a_0, a_1, \dots, a_m) = \sum_{i=1}^N (a_0 + a_1 x_i + a_2 x_i^2 + \dots + a_m x_i^m - f_i)^2$$

Приравнивая эти выражения нулю и собирая коэффициенты при неизвестных, получим следующую систему линейных уравнений:

$$Na_0 + a_1 \sum_{i=1}^N x_i + a_2 \sum_{i=1}^N x_i^2 + \dots + a_m \sum_{i=1}^N x_i^m = \sum_{i=1}^N f_i$$

$$a_0 \sum_{i=1}^N x_i + a_1 \sum_{i=1}^N x_i^2 + a_2 \sum_{i=1}^N x_i^3 + \dots + a_m \sum_{i=1}^N x_i^{m+1} = \sum_{i=1}^N x_i f_i$$

.....

$$a_0 \sum_{i=1}^N x_i^m + a_1 \sum_{i=1}^N x_i^{m+1} + a_2 \sum_{i=1}^N x_i^{m+2} + \dots + a_m \sum_{i=1}^N x_i^{2m} = \sum_{i=1}^N x_i^m f_i$$

В случае полинома первого порядка $m=1$, т.е. $\varphi(x) = a_0 + a_1 x$, система нормальных уравнений примет вид:

$$Na_0 + a_1 \sum_{i=1}^N x_i = \sum_{i=1}^N f_i$$

$$a_0 \sum_{i=1}^N x_i + a_1 \sum_{i=1}^N x_i^2 = \sum_{i=1}^N x_i f_i$$

При $m=2$ имеем:

$$Na_0 + a_1 \sum_{i=1}^N x_i + a_2 \sum_{i=1}^N x_i^2 = \sum_{i=1}^N f_i$$

$$a_0 \sum_{i=1}^N x_i + a_1 \sum_{i=1}^N x_i^2 + a_2 \sum_{i=1}^N x_i^3 = \sum_{i=1}^N x_i f_i$$

$$a_0 \sum_{i=1}^N x_i^2 + a_1 \sum_{i=1}^N x_i^3 + a_2 \sum_{i=1}^N x_i^4 = \sum_{i=1}^N x_i^2 f_i$$

Для проверки согласия построенной кривой регрессии с результатами эксперимента обычно вводят следующие числовые характеристики:

коэффициент корреляции (линейная зависимость), корреляционное отношение и коэффициент детерминированности.

$$r^2 = 1 - \frac{S_{ост}}{S_{полн}}$$

Вычисление коэффициента детерминированности:

$$\sum_{i=1}^n (y_i - y_i^T)^2$$

- остаточная сумма квадратов, характеризующая отклонение экспериментальных данных от теоретических.

$$\sum_{i=1}^n (y_i - \bar{y})^2$$

- Сполн - полная сумма квадратов, где \bar{y} - среднее значение y_i .

Пример.

Экспериментальные данные о значениях переменных x и y приведены в таблице.

x	y
0,25	5,8945
0,75	5,7276
1,25	5,6045
1,75	5,4955
2,25	5,3637
2,75	5,2422
3,25	5,1196
3,75	5,0045

Используя метод наименьших квадратов, аппроксимировать эти данные линейной зависимостью $y=ax+b$ (найти параметры a и b).

Нанести точки x и y на точечную диаграмму и с помощью команды Добавить линию тренда... построить линейную линию тренда и показать уравнение на диаграмме. Сравнить полученные коэффициенты a и b.

Решение:

Заносим на лист MS Excel исходные данные и вычисляем необходимые суммы:

	A	B	C	D
1	x	y	x^2	xy
2	0,25	5,8945	0,0625	1,4736
3	0,75	5,7276	0,5625	4,2957
4	1,25	5,6045	1,5625	7,0056
5	1,75	5,4955	3,0625	9,6171
6	2,25	5,3637	5,0625	12,068
7	2,75	5,2422	7,5625	14,416
8	3,25	5,1196	10,563	16,639
9	3,75	5,0045	14,063	18,767
10	16	43,452	42,5	84,282

Рис. 1 Фрагмент листа MS Excel в режиме отображения данных

	A	B	C	D
1	x	y	x^2	xy
2	0,25	5,8945	=A2*A2	=A2*B2
3	0,75	5,7276	=A3*A3	=A3*B3
4	1,25	5,6045	=A4*A4	=A4*B4
5	1,75	5,4955	=A5*A5	=A5*B5
6	2,25	5,3637	=A6*A6	=A6*B6
7	2,75	5,2422	=A7*A7	=A7*B7
8	3,25	5,1196	=A8*A8	=A8*B8
9	3,75	5,0045	=A9*A9	=A9*B9
10	=СУММ(A2:A9)	=СУММ(B2:B9)	=СУММ(C2:C9)	=СУММ(D2:D9)

Рис. 2 Фрагмент листа MS Excel в режиме отображения формул

Выписываем коэффициенты при неизвестных и столбец свободных членов.

Решаем СЛАУ матричным методом:

F	G	H	I
	A		B
42,5	16		84,282
16	8		43,452
	A^{-1}		
0,0952	-0,19	a=	-0,250
-0,19	0,506	b=	5,931

Рис. 3 Фрагмент листа MS Excel в режиме отображения данных

F	G	H	I
	A		B
=C10	=A10		=D10
=A10	=8		=B10
	A ⁻¹		
=МОБР(F4:G5)	=МОБР(F4:G5)	a=	=МУМНОЖ(F8:G9;I4:I5)
=МОБР(F4:G5)	=МОБР(F4:G5)	b=	=МУМНОЖ(F8:G9;I4:I5)

Рис. 4 Фрагмент листа MS Excel в режиме отображения данных

Наносим наши точки на точечную диаграмму и строим линию тренда:

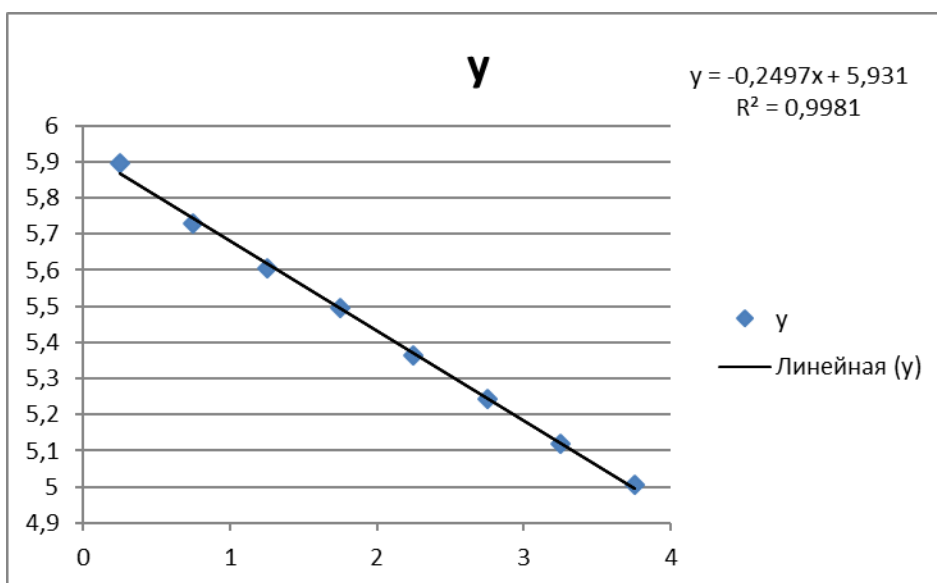


Рис. 5 Диаграмма с линией тренда

Вывод: Коэффициенты a и b совпадают, следовательно, вычисления проведены верно.

Задание для самостоятельной работы:

Экспериментальные данные о значениях переменных x и y приведены в таблице.

Используя метод наименьших квадратов, аппроксимировать эти данные линейной зависимостью $y = a_0 + a_1x$ (найти параметры a_0 и a_1) и квадратичной зависимостью $a_0 + a_1x + a_2x^2$ (найти a_0 , a_1 и a_2). Вычислить r^2 для каждой аппроксимирующей функции.

Нанести точки x и y на точечную диаграмму и с помощью команды Добавить линию тренда... построить линейную и полиномиальную линии тренда и показать уравнение на диаграмме. Сравнить полученные коэффициенты и r^2 .

Пример оформления:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	x	y	x ²	xy	y1=a ₀ +a ₁ x	x ³	x ⁴	x ² y	y2=a ₀ +a ₁ x+a ₂ x ²	(y-y1) ²	(y-y2) ²	(y-y _{ср}) ²	
2	-5,0	0,5	25,00	-2,50	0,89	-125,00	625,00	12,50	0,61	0,152	0,011	0,667	
3	-3,5	1,2	12,25	-4,20	1,02	-42,88	150,06	14,70	1,05	0,032	0,023	0,014	
4	-2,0	1,4	4,00	-2,80	1,15	-8,00	16,00	5,60	1,38	0,061	0,000	0,007	
5	1,5	1,6	2,25	2,40	1,46	3,38	5,06	3,60	1,72	0,020	0,015	0,080	
6	3,3	1,7	10,56	5,53	1,61	34,33	111,57	17,96	1,67	0,008	0,001	0,147	
7	5,0	1,5	25,00	7,50	1,77	125,00	625,00	37,50	1,47	0,070	0,001	0,034	
8	-0,8	7,9	79,06	5,93	7,90	-13,17	1532,69	91,86	7,90	0,3433	0,0520	0,9483	Суммы
9													
10							Y _{ср}	\bar{y}	1,3				$\sum_{i=1}^n (y_i - \bar{y})^2$
11										r ² для y1	r ² для y2		
12										0,638049	0,945143		
13		A1			B1								
14		6	-0,8		7,9								
15		-0,8	79,1		5,9								
16		a0	1,328										
17		a1	0,088										
18													
19		A2				B2							
20		6	-0,8	79,1		7,9							
21		-0,8	79,1	-13,2		5,9							
22		79,1	-13,2	1532,7		91,9							
23													
24		a0	1,648										
25		a1	0,087										
26		a2	-0,024										

Рис. 6 Фрагмент листа MS Excel с примерным решением задачи

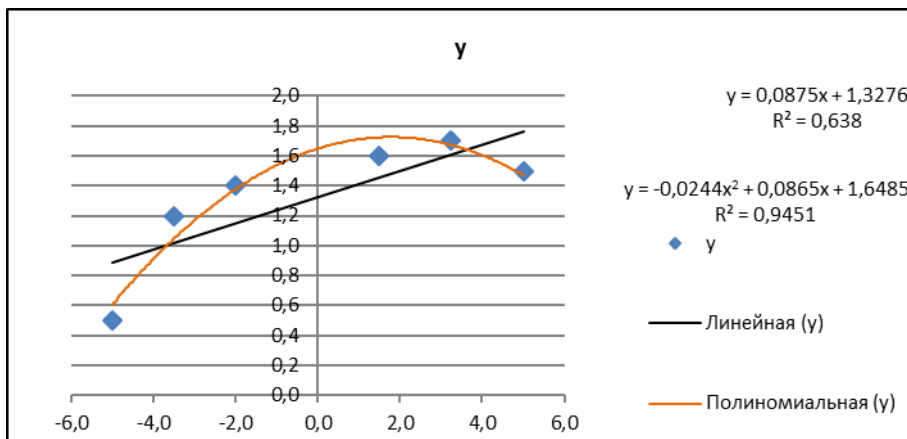


Рис. 7 Диаграмма с линиями тренда

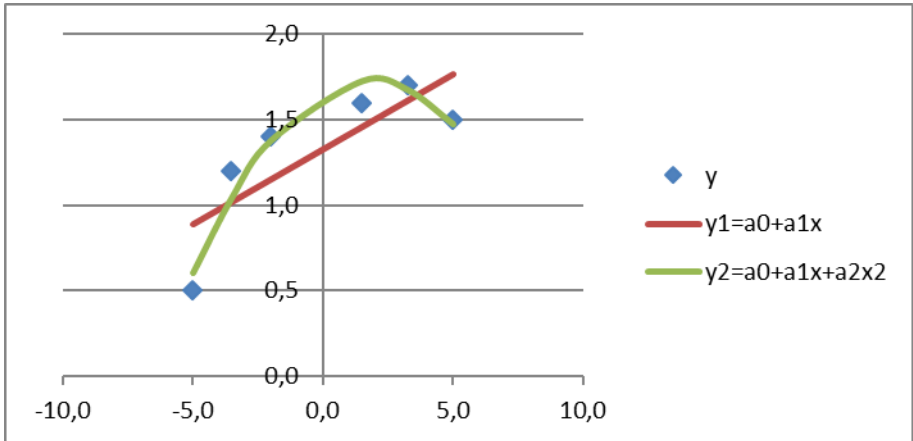


Рис. 8 Диаграмма, построенная по данным расчетов

Варианты заданий:

ВАРИАНТ 1													
X_i	0,08	0,56	1,23	1,89	2,22	2,78	3,43	4,01	4,67	5,08	5,59	6,09	
Y_i	0,05	3,65	8,81	15,09	17,87	22,21	23,54	31,21	35,32	39,45	43,39	47,89	
X_i	6,44	7,05	7,34	7,91	8,05	8,45	9,12	9,54	10,32	10,43	11,34	11,93	12,43
Y_i	51,23	55,52	58,45	62,48	63,23	68,56	71,67	75,56	79,09	79,98	87,09	94,23	95,56
ВАРИАНТ 2													
X_i	2,54	4,43	5,32	5,99	6,12	6,45	7,32	7,76	8,23	8,87	9,08	9,66	
Y_i	1,12	2,29	2,99	3,09	3,78	3,15	4,44	4,12	4,11	4,67	4,59	4,88	
X_i	10,22	10,88	11,45	12,67	14,02	14,87	15,55	16,71	17,77	18,98	20,34	22,45	23,49
Y_i	5,19	5,66	6,01	6,33	7,33	7,54	8,87	8,35	9,06	9,76	10,34	12,43	12,21
ВАРИАНТ 3													
X_i	0,81	1,23	2,12	2,97	3,56	4,49	5,27	5,94	6,56	6,89	6,95	7,21	
Y_i	20,65	29,56	35,67	39,45	45,78	53,98	74,87	98,99	104,8	110,8	115,8	121,6	
X_i	7,45	7,99	8,56	8,92	9,34	9,76	9,98	10,45	10,97	10,76	11,65	11,95	12,56
Y_i	126,8	120,9	115,9	112,9	102,7	95,89	70,67	55,89	45,34	40,67	30,78	28,78	15,46
ВАРИАНТ 4													
X_i	0,21	1,19	2,43	3,12	4,54	4,98	5,49	6,07	6,81	7,21	7,96	8,32	
Y_i	1,62	8,65	16,76	24,45	32,87	40,09	43,56	48,45	52,21	57,34	63,31	67,45	
X_i	9,43	10,21	11,54	12,33	13,21	14,72	15,53	16,23	17,32	18,43	19,38	20,45	21,22
Y_i	72,87	81,34	89,45	97,77	105,3	112,6	121,9	108,5	126,5	144,3	160,5	161,3	170,6
ВАРИАНТ 5													
X_i	0,44	0,72	1,26	1,58	1,87	2,12	2,65	3,12	3,39	3,81	4,21	4,59	
Y_i	1,54	2,32	3,86	7,21	11,11	15,43	21,54	29,54	39,67	53,65	89,76	95,65	
X_i	4,79	5,04	5,41	5,89	6,23	6,72	6,99	7,05	7,34	7,69	8,23	8,83	9,21
Y_i	100,7	91,09	84,43	50,06	43,76	31,34	22,12	16,23	11,11	7,78	4,87	2,99	1,91
ВАРИАНТ 6													

Xi	2,85	3,45	4,23	4,77	5,46	6,85	7,32	7,79	8,34	8,87	9,41	9,79	
Yi	300,8	234,9	188,8	132,2	98,97	82,45	76,45	34,64	25,45	19,76	14,29	13,62	
Xi	10,54	10,84	12,45	12,89	13,62	14,51	15,21	15,84	16,32	16,65	17,34	18,53	19,32
Yi	14,51	15,21	15,84	21,34	31,43	40,78	81,12	92,34	107,5	130,5	185,8	243,7	287,8
ВАРИАНТ 7													
Xi	0,76	1,23	3,87	5,43	5,94	7,12	7,97	8,55	9,65	10,78	11,54	12,23	
Yi	4,67	8,56	25,78	38,65	40,76	48,54	55,89	57,76	64,86	71,45	78,98	85,91	
Xi	13,86	14,53	15,48	16,75	17,45	18,81	19,64	20,35	21,45	22,23	23,45	24,67	25,78
Yi	91,25	100,6	105,4	112,3	119,1	125,9	135,5	140,8	149,4	154,5	161,5	168,5	180,7
ВАРИАНТ 8													
Xi	0,55	1,43	2,87	3,54	4,63	5,34	6,35	6,86	7,58	8,23	9,45	9,97	
Yi	1,54	3,78	6,78	9,76	17,87	47,87	98,09	154,1	198,1	246	287,7	299,9	
Xi	10,55	11,34	12,12	13,56	14,65	14,76	15,86	16,33	17,21	17,69	18,54	19,08	19,78
Yi	357	322	296,9	255,9	200,7	167,4	107,7	54,98	20,76	11,78	7,87	4,08	2,78
ВАРИАНТ 9													
Xi	0,09	0,15	0,87	1,85	2,75	2,85	3,08	3,56	3,92	4,07	4,44	4,76	
Yi	2,54	2,76	2,98	6,45	20,65	22,47	25,41	54,89	80,87	85,75	156,9	198	
Xi	5,05	5,45	5,72	5,91	6,98	7,12	7,65	8,13	8,54	8,86	9,08	9,45	9,55
Yi	242,7	345,9	400,9	456,8	601	645,7	789,1	801	867,7	900,5	956,9	977,1	999,5
ВАРИАНТ 10													
Xi	2,23	2,87	3,66	3,97	4,86	5,34	5,76	6,23	7,45	8,76	9,65	10,76	
Yi	8,09	12,76	15,76	16,76	20,97	22,21	23,78	25,55	30,97	32,21	39,54	40,76	
Xi	11,65	12,45	16,76	17,61	18,05	18,87	19,97	20,85	22,21	24,65	25,32	26,43	27,65
Yi	45,65	46,78	64,87	70,76	75,45	76,98	82,07	83,09	88,07	96,97	104,4	112,9	120,1
ВАРИАНТ 11													
Xi	1,08	2,21	3,87	4,07	4,76	5,14	5,98	6,32	6,99	7,43	7,54	7,84	
Yi	200,5	149,4	129	116,6	87,76	60,98	40,98	29,98	24,65	19,99	15,67	13,98	
Xi	8,07	8,57	8,94	9,32	9,86	10,05	10,54	10,87	11,23	11,46	11,96	12,54	13,86
Yi	12,09	14,97	16,09	18,65	22,09	31,65	38,56	57,98	85,98	114,4	130,9	151	196,8
ВАРИАНТ 12													
Xi	1,05	1,65	2,08	2,76	2,99	3,65	4,05	4,15	4,39	4,76	5,08	5,43	
Yi	3,44	6,76	9,08	17,98	27,78	40,43	53,87	59,96	70,08	85,96	95,06	101	
Xi	5,89	6,43	6,91	7,13	7,34	8,01	8,54	9,01	9,54	9,85	10,06	10,42	10,89
Yi	121,8	112,8	99,05	87,95	72,08	60,87	55,08	44,41	25,97	18,64	11,43	8,87	5,51
ВАРИАНТ 13													
Xi	0,66	1,34	3,87	5,65	6,99	8,98	10,45	11,88	13,78	15,02	16,22	17,11	
Yi	3,87	6,78	17,56	26,65	37,87	44,54	51,87	59,45	70,56	72,86	82,43	85,56	
Xi	18,43	19,56	20,32	20,53	20,97	21,34	21,98	22,41	23,05	23,54	24,29	24,79	25,43
Yi	90,45	92,34	96,85	100,8	104,3	106,5	109,8	111,9	116,7	117,7	120,5	121,8	128,8
ВАРИАНТ 14													
Xi	0,86	1,54	2,78	3,99	5,34	5,78	6,45	7,23	7,65	7,91	8,34	8,78	
Yi	1,89	6,76	9,79	15,98	54,67	76,98	122,3	167,9	188,7	219,9	256,8	277,7	

Xi	8,88	9,15	9,25	9,54	9,75	10,32	10,67	10,78	11,22	11,39	11,39	11,43	12,19
Yi	300,5	342,6	378,5	399,9	426	466,9	489,9	500,1	555,4	579	623,4	630,8	686
	ВАРИАНТ 15												
Xi	3,47	3,75	4,05	4,51	4,55	5,05	5,13	5,35	5,46	5,79	6,17	6,58	
Yi	3,44	5,61	8,21	7,28	14,69	14,99	15,43	12,4	16,94	17,49	19,31	21,5	
Xi	6,6	6,81	7,26	7,77	8,04	8,38	8,68	8,8	9,11	9,13	9,38	9,63	10,06
Yi	20,09	21,71	22,32	23,83	22,26	28,14	23,02	21,46	30,09	24,34	22,54	26,25	24,9
	ВАРИАНТ 16												
Xi	1,04	1,17	1,22	1,33	1,47	1,56	1,72	1,88	1,93	1,94	1,94	2,08	
Yi	-6,4	-5,87	-4,27	1,15	4,33	3,46	-1,78	-2,61	-0,53	2,88	1,45	0,83	
Xi	2,25	2,25	2,32	2,37	2,57	2,69	2,77	2,87	2,99	3,07	3,17	3,22	3,33
Yi	5,45	6,21	9,09	0,95	5,39	13,8	14,22	16,88	15,73	12,12	15,85	18,64	24,6