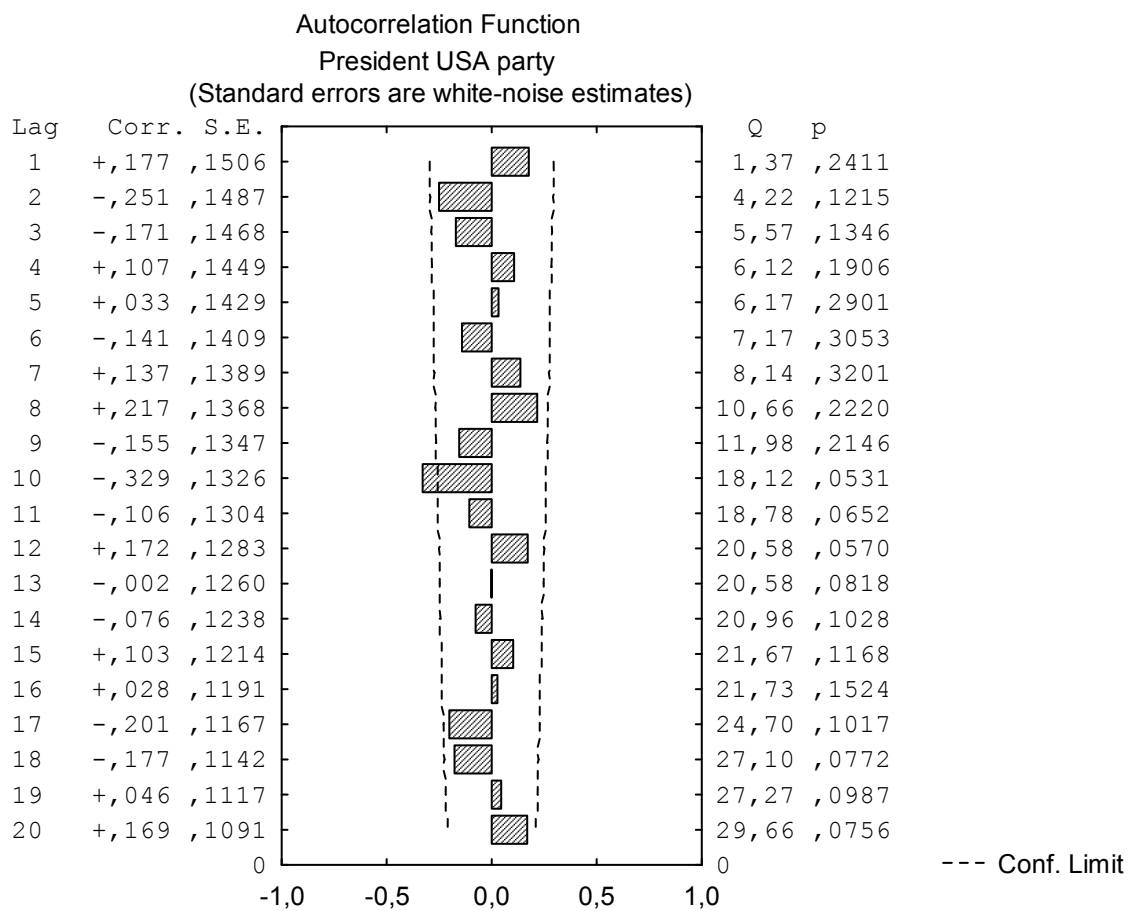


РЕЗУЛЬТАТЫ ВЫЧИСЛЕНИЙ

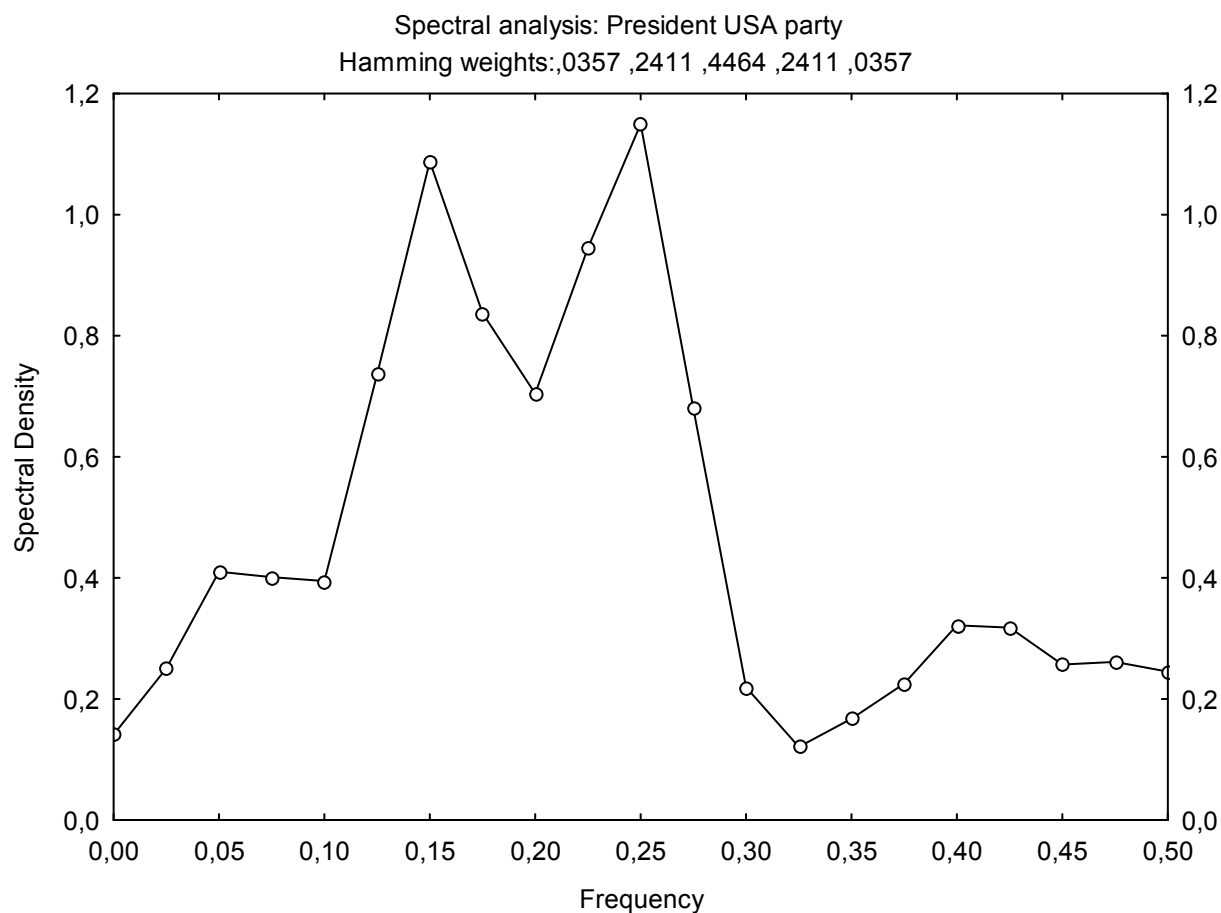
К статье А.А.Давыдова «Партийная принадлежность
Президентов США, 1852-2016: анализ и прогнозирование»

Рисунок 1

**Значения автокорреляционной функции последовательности партийной
принадлежности Президентов США, 1852-2012 гг.**



Спектральный анализ (Фурье анализ) последовательности партийной принадлежности Президентов США, 1852-2012 гг.



Примечания. Наблюдаются циклы продолжительностью 24 и 40 лет.

«Ксения». В «Ксении» реализованы переборные алгоритмы выявления закономерностей в дискретных последовательностях, обобщение множества частных прогнозов и эвристические методы экспертного оценивания.

Результаты прогнозирования

Исходная последовательность партийной принадлежности Президентов США за период 1852-2012 гг.:

DDRRRRRRDRDRRRRRDDRRRRDDDDDRRDRRRDRRRDDRRDD

где R – республиканец, D – демократ

Частота встречаемости элементов последовательности

D – 44%

R – 56%

Вероятность перехода соседних элементов в последовательности:

	D	R
D	9	8
R	8	15

	D	R
D	0.529	0.471
R	0.348	0.652

Вероятность перехода двух соседних элементов в третий:

	D	R
DD	3	5
DR	1	7
RR	7	8
RD	5	3

	D	R
DD	0.375	0.625
DR	0.125	0.875
RR	0.467	0.533
RD	0.625	0.375

Пояснение. После победы Б.Обамы – демократа в 2008 и 2012 гг., т.е. в записи DD, с вероятностью 0.625 на следующих выборах побеждает республиканец – Д.Трамп.

Вероятность перехода трех соседних элементов в четвертый:

	D	R
DDR	0	5
DRR	3	4
RRR	4	4
RRD	5	2
RDR	1	2
DRD	0	1
RDD	1	3
DDD	2	1

Пояснение. После RDD (за период 2004–2012 гг.) с вероятностью 0.75 следует R, т.е. побеждает республиканец Д.Трамп.

Вероятность перехода четырех соседних элементов в пятый:

	D	R
DDRR	3	2
DRRR	2	2
RRRR	2	2
RRRD	3	1
RRDR	1	1
RDRD	0	1
DRDR	0	1
RDRR	0	2
RRDD	1	3
RDDR	0	3
RDDD	1	0
DDDD	1	1
DDDR	0	1
DRRD	2	1

Пояснение. После RRDD (за период 2000-2012 гг.) с вероятностью 0.75 следует R, т.е. побеждает республиканец Д.Трамп.

Прогноз одного элемента:

По общей частоте встречаемости: прогноз - R.

По вероятности переходов соседних элементов: прогноз - D.

По вероятности переходов двух соседних элементов: прогноз - R.

По вероятности переходов трех соседних элементов: прогноз - R.

По вероятности переходов четырех соседних элементов: прогноз - R.

По генеральной совокупности наиболее вероятен прогноз - R.

Частота встречаемости двух соседних элементов последовательности:

	Кол-во раз	Доля, %
DD	9	22%
DR	8	20%
RR	15	37%
RD	8	20%

Вероятность перехода двух соседних элементов в два соседних:

	DD	DR	RR	RD
DD	2	1	5	0
DR	0	1	4	3
RR	5	2	4	4
RD	1	3	2	1

Вероятность перехода трех соседних элементов в два следующих:

	DD	DR	RR	RD
DDR	0	0	2	3
DRR	2	1	2	2
RRR	3	1	2	2
RRD	1	3	1	1
RDR	0	1	2	0
DRD	0	0	1	0
RDD	1	0	3	0
DDD	1	1	1	0

Прогноз двух элементов:

По общей частоте встречаемости: прогноз – RR.

По вероятности переходов двух элементов: прогноз – DD.

По вероятности переходов трех элементов: прогноз – DD, RR, RD.

Циклы последовательности серий:

N серии	Серия	Длина цикла
1	DD	нет периодичности
2	RRRRRR	нет периодичности
3	D	нет периодичности
4	R	нет периодичности
5	RRRR	нет периодичности
6	RRR	5
7	DDDDD	нет периодичности
8	RR	нет периодичности

Номер серии	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Элемент серии	D	R	D	R	D	R	D	R	D	R	D	R	D	R	D	R	D
Длина серии	2	6	1	1	1	4	2	3	5	2	2	2	1	3	2	2	2

Средняя длина серии = 2.4

Средняя длина серии D = 2.1

Средняя длина серии R = 2.7

Таблица переходов серий:

	D	R
D	0	8
R	8	0

Таблица переходов длин серий:

	2	6	1	4	3	5
2	4	1	1	0	1	0
6	0	0	1	0	0	0
1	0	0	2	1	1	0
4	1	0	0	0	0	0
3	1	0	0	0	0	1
5	1	0	0	0	0	0

Прогноз серии:

Наиболее вероятна следующая серия: R.

Наиболее вероятна следующая длина серии: 2.

Циклы последовательности элементов:

№ элемента	Обозначение	Длина цикла
1	D	0, 6, 1, 4, 0, 3, 0, 0, 0, 0, 2, 0, 2, 3, 0, 2, 0
2	D	6, 1, 4, 0, 3, 0, 0, 0, 0, 2, 0, 2, 3, 0, 2, 0
3	R	0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 2, 0, 0, 5, 0, 2, 0, 1, 0, 0, 2, 0
4	R	0, 0, 0, 0, 1, 1, 0, 0, 0, 2, 0, 0, 5, 0, 2, 0, 1, 0, 0, 2, 0
5	R	0, 0, 0, 1, 1, 0, 0, 0, 2, 0, 0, 5, 0, 2, 0, 1, 0, 0, 2, 0
6	R	0, 0, 1, 1, 0, 0, 0, 2, 0, 0, 5, 0, 2, 0, 1, 0, 0, 2, 0
7	R	0, 1, 1, 0, 0, 0, 2, 0, 0, 5, 0, 2, 0, 1, 0, 0, 2, 0
8	R	1, 1, 0, 0, 0, 2, 0, 0, 5, 0, 2, 0, 1, 0, 0, 2, 0
9	D	1, 4, 0, 3, 0, 0, 0, 0, 2, 0, 2, 3, 0, 2, 0
10	R	1, 0, 0, 0, 2, 0, 0, 5, 0, 2, 0, 1, 0, 0, 2, 0
11	D	4, 0, 3, 0, 0, 0, 0, 2, 0, 2, 3, 0, 2, 0
12	R	0, 0, 0, 2, 0, 0, 5, 0, 2, 0, 1, 0, 0, 2, 0
13	R	0, 0, 2, 0, 0, 5, 0, 2, 0, 1, 0, 0, 2, 0
14	R	0, 2, 0, 0, 5, 0, 2, 0, 1, 0, 0, 2, 0
15	R	2, 0, 0, 5, 0, 2, 0, 1, 0, 0, 2, 0
16	D	0, 3, 0, 0, 0, 0, 2, 0, 2, 3, 0, 2, 0
17	D	3, 0, 0, 0, 0, 2, 0, 2, 3, 0, 2, 0
18	R	0, 0, 5, 0, 2, 0, 1, 0, 0, 2, 0
19	R	0, 5, 0, 2, 0, 1, 0, 0, 2, 0
20	R	5, 0, 2, 0, 1, 0, 0, 2, 0
21	D	0, 0, 0, 0, 2, 0, 2, 3, 0, 2, 0
22	D	0, 0, 0, 2, 0, 2, 3, 0, 2, 0
23	D	0, 0, 2, 0, 2, 3, 0, 2, 0
24	D	0, 2, 0, 2, 3, 0, 2, 0
25	D	2, 0, 2, 3, 0, 2, 0
26	R	0, 2, 0, 1, 0, 0, 2, 0
27	R	2, 0, 1, 0, 0, 2, 0
28	D	0, 2, 3, 0, 2, 0
29	D	2, 3, 0, 2, 0
30	R	0, 1, 0, 0, 2, 0
31	R	1, 0, 0, 2, 0
32	D	3, 0, 2, 0
33	R	0, 0, 2, 0
34	R	0, 2, 0
35	R	2, 0
36	D	0, 2, 0
37	D	2, 0
38	R	0
39	R	-
40	D	0
41	D	-

ОКОНЧАТЕЛЬНЫЙ ПРОГНОЗ:

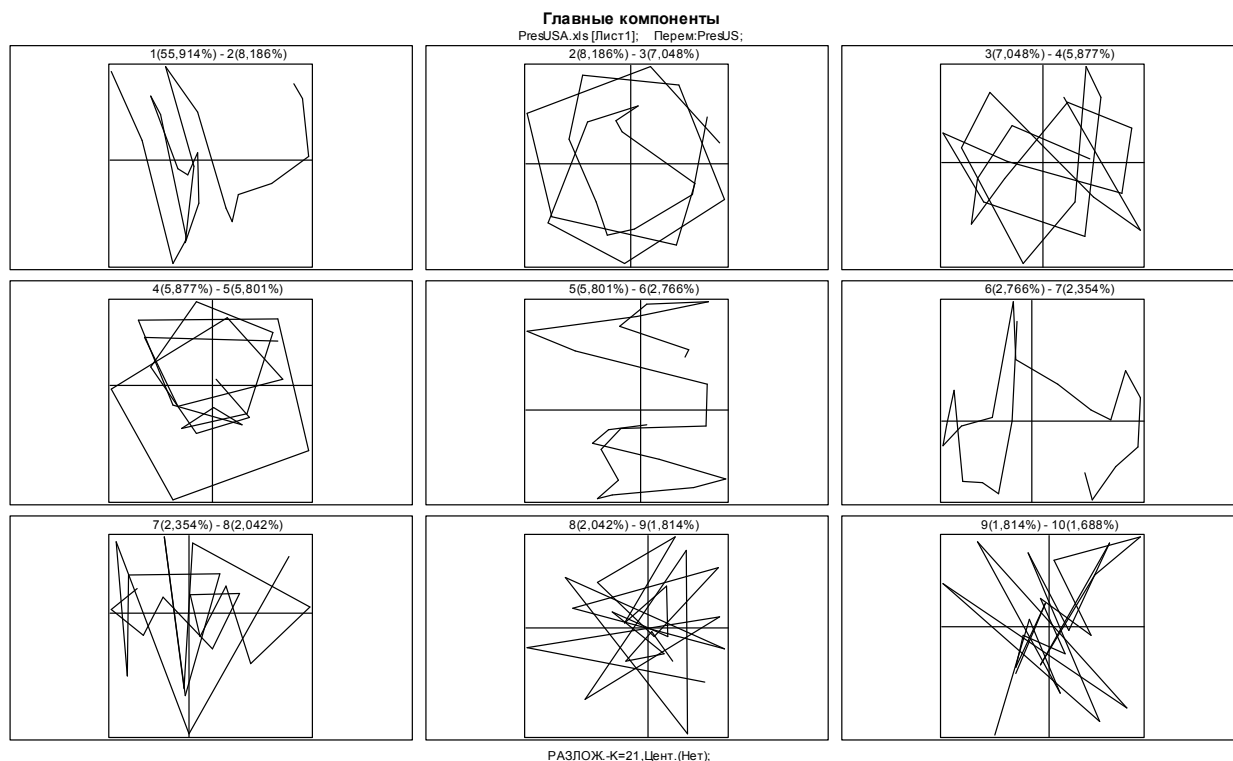
Исходная последовательность партийной принадлежности Президентов США за период 1852–2012 гг.: D D R R R R R R R D R D R R R R R D D R R R D D D D R R D D R R R R R D D R R D D.

Прогноз следующего элемента последовательности: R, где R – республиканец. Прогноз корректный, поскольку Д.Трамп – республиканец.

Прогноз SVD (Singular Value Decomposition). В SVD используется метод главных компонент при разложении матрицы, составленной из членов последовательности. Длина «окна» – 21 временная точка.

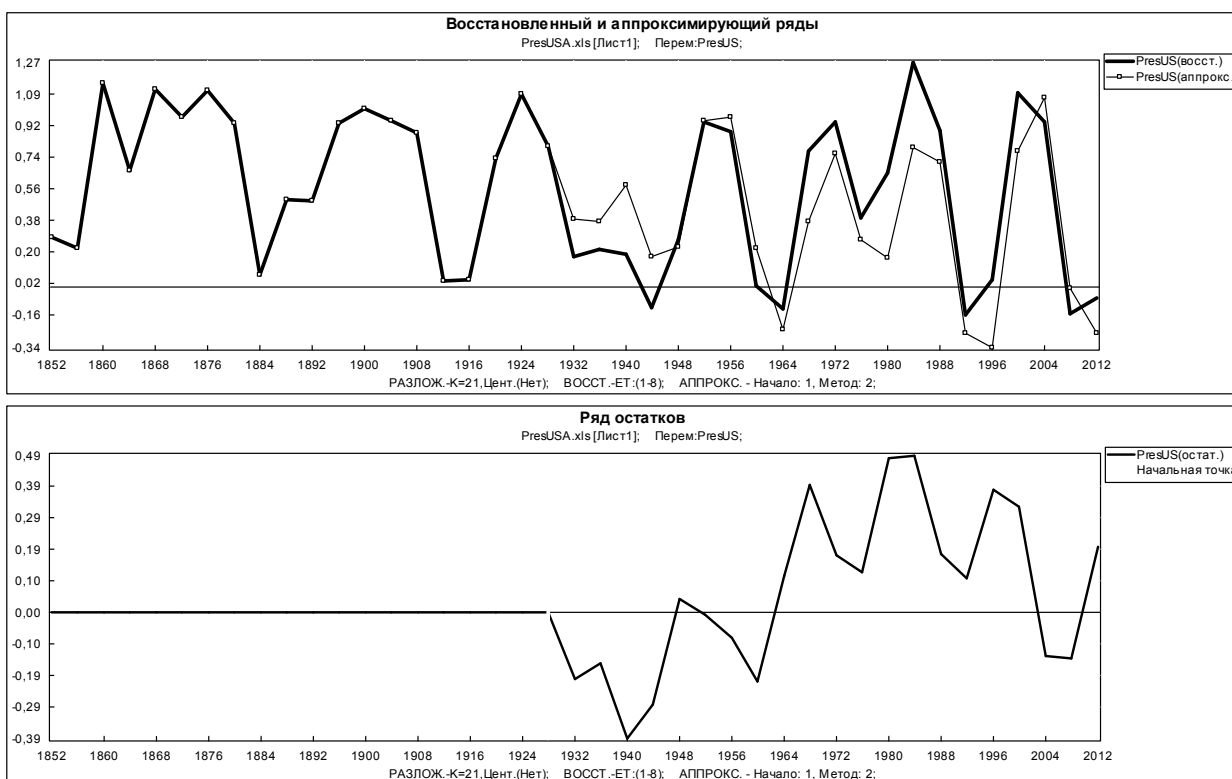
Рисунок 3

Главные компоненты последовательности партийной принадлежности Президентов США



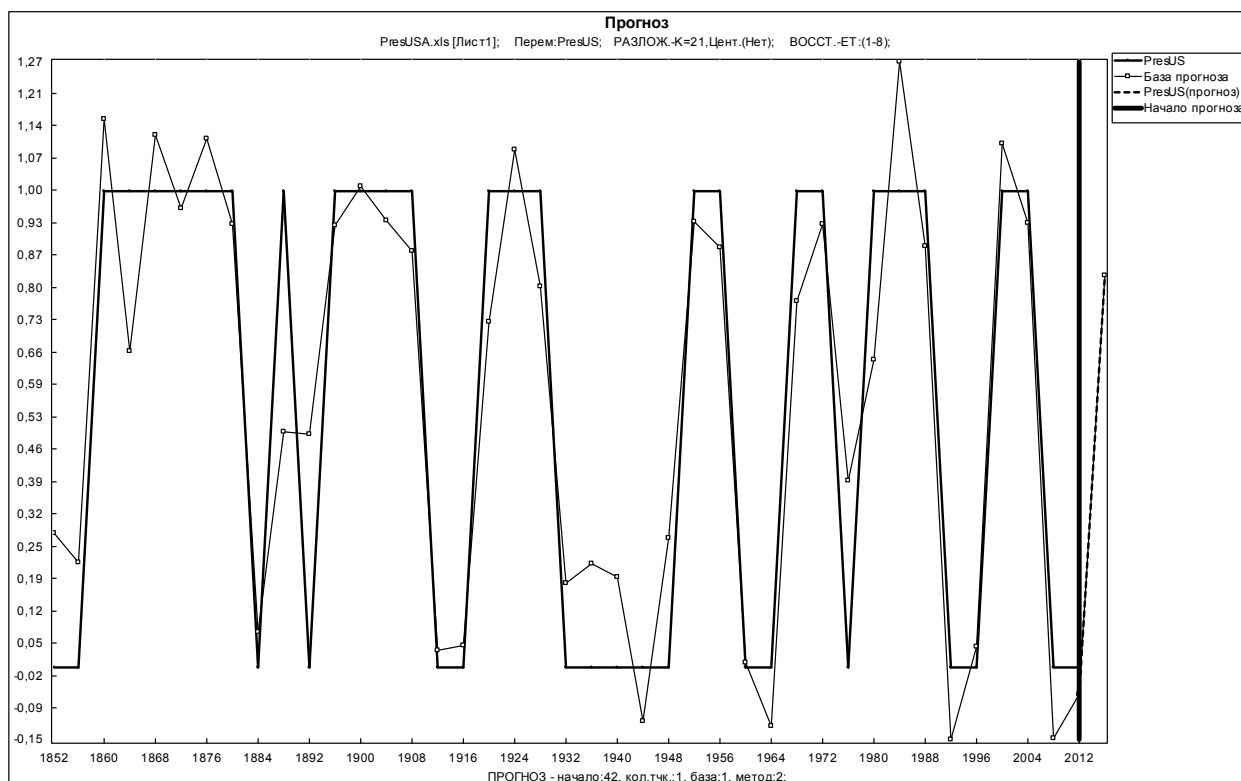
Для восстановления и аппроксимации исходной последовательности были использованы пять выявленных в ходе проведенного анализа компонент, которые объясняют 82,8% суммарной дисперсии. Первая компонента – 55,91%, вторая компонента – 8,19%, третья компонента – 7,05%, четвертая компонента – 5,88%, пятая компонента – 5,8%. На рисунке 4 представлены восстановленная и аппроксимирующая последовательности по результатам SVD.

**Восстановленная и аппроксимирующая последовательности
партийной принадлежности Президентов США**



На рисунке 5 представлен прогноз бинарной последовательности для 2016 г. Базовый ряд прогноза – восстановленный, метод прогнозирования рекуррентный, на одну временную точку вперед.

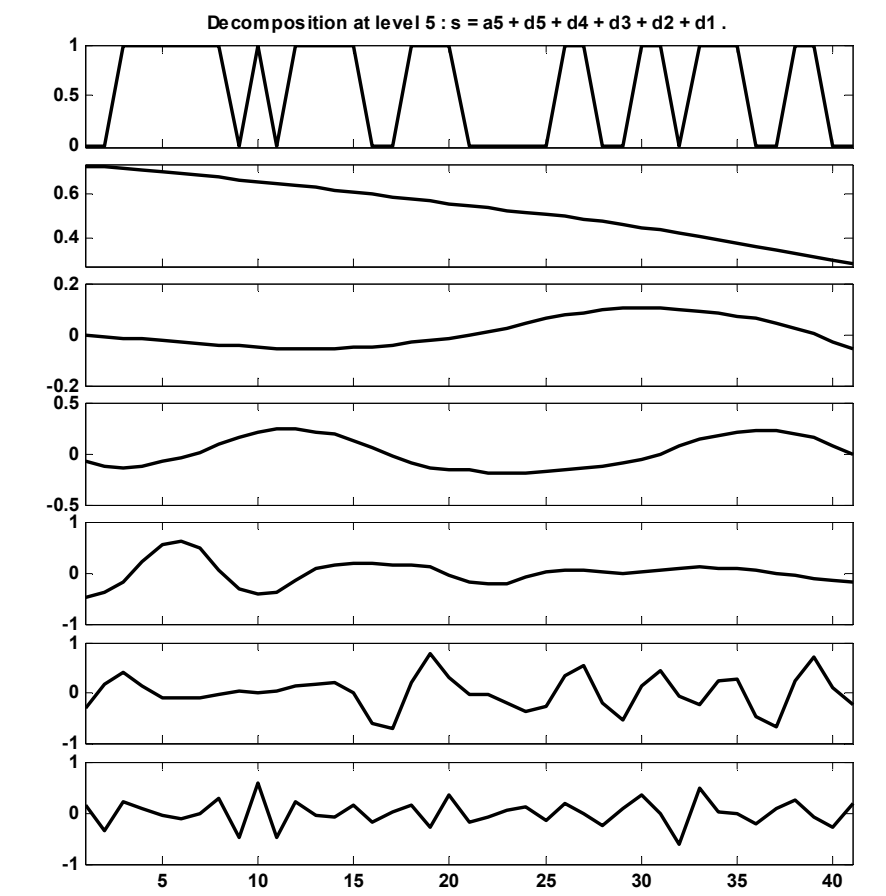
Прогноз партийной принадлежности Президента США на 2016 г.



Прогноз, представленный на рисунке 5, корректно предсказывает победу республиканца – Д.Трампа в 2016 г.

Вейвлет-анализ. Прогнозирование осуществлялось на основе вейвлет-преобразования с использованием вейвлета Добеши из пакета Wavelet Toolbox MATLAB [MATLAB]. Вейвлет-анализ позволяет наглядно представить суперпозицию различных частотных колебаний во времени и тренд – долговременную тенденцию изменения анализируемой последовательности. На рисунке 6 представлена вейвлет-декомпозиция анализируемой бинарной последовательности, выполненная с помощью вейвлет-преобразования Добеши с параметром порядка 5, на 5 уровнях декомпозиции.

**Вейвлет-декомпозиция последовательности партийной принадлежности
Президентов США, 1852-2012 гг.**



Примечания. По оси x отложены порядковые номера членов последовательности. Первый сверху фрагмент рисунка – анализируемая бинарная последовательность, второй сверху фрагмент рисунка – тренд, третий и четвертый фрагменты рисунка – циклы.

Прогноз на 2012 и 2016 гг. был реализован в пакете Wavelet Toolbox MATLAB с использованием процедуры Signal Extension на основе SWT преобразования (вейвлет Добеши, значение параметра порядка равно 5, уровней декомпозиции сигнала -5).

«Нейронная» сеть. Ниже представлены результаты обучения, полученные с помощью «нейронной» сети класса Support Vector Machine (SVM), реализованной в пакете NeuroSolutions [NeuroSolutions]. Метод опорных векторов (SVM) - перевод исходных векторов в пространство более высокой размерности и поиск разделяющей гиперплоскости с максимальным зазором в данном пространстве.

В качестве «входа» (независимой переменной) выступали календарные даты с 1852 по 2012 гг., а в качестве «выхода» (зависимой переменной) выступали значения бинарной последовательности. Была проведена серия вычислительных экспериментов с различными классами «нейронных» сетей, различной архитектурой сетей, методами обучения, передаточными функциями и т.д., реализованными в пакете NeuroSolutions [NeuroSolutions]. Обучение «нейронной» сети осуществлялось за период с 1852 по 2012 гг., в течение 30000 эпох. В таблице 1 представлены полученные результаты аппроксимации исходной бинарной последовательности с помощью «нейронной сети» класса Support Vector Machine (SVM).

Таблица 1

Результаты обучения «нейронной» сети Support Vector Machine (SVM)

Ошибка	Демократ (0)	Республиканец (1)
Средняя квадратичная ошибка (MSE)	0,019860729	0,024418732
Средняя абсолютная ошибка (MAE)	0,120911541	0,141334215
Минимальная абсолютная ошибка	0,030786439	0,054908639
Максимальная абсолютная ошибка	0,289369484	0,316121119
Точность аппроксимации, %	100	100

Логит, Пробит, Тобит модели

Модель 1: Логит, использованы наблюдения 11-41 (T = 31)
Зависимая переменная: PresUS_1_0

	Коэффициент	Ст. ошибка	z	P-значение
const	15,1589	36,4678	0,4157	0,67764
Year	-0,0068024	0,0179884	-0,3782	0,70532
PresUS_1_0_1	0,596909	1,29825	0,4598	0,64567
PresUS_1_0_2	-1,39796	1,16885	-1,1960	0,23169
PresUS_1_0_3	-1,2647	1,13457	-1,1147	0,26498
PresUS_1_0_4	0,5736	1,19111	0,4816	0,63011
PresUS_1_0_5	-0,139758	1,12348	-0,1244	0,90100
PresUS_1_0_6	-1,28434	1,16595	-1,1015	0,27066
PresUS_1_0_7	1,15234	1,20783	0,9541	0,34006
PresUS_1_0_8	1,33501	1,15315	1,1577	0,24698
PresUS_1_0_9	-1,3549	1,27224	-1,0650	0,28689
PresUS_1_0_10	-1,3303	1,13221	-1,1750	0,24001
Среднее зав. перемен	0,516129	Ст. откл. зав. перемен		0,248418
R-квадрат Макфаддена	0,333907	Испр. R-квадрат		-0,224975
Лог. правдоподобие	-14,30196	Крит. Акаике		52,60392
Крит. Шварца	69,81177	Крит. Хеннана-Куинна		58,21325

Количество 'корректно предсказанных' случаев = 24 (77,4%)
 $f(\beta'x)$ для среднего значения независимых переменных = 0,248
 Критерий отношения правдоподобия: Хи-квадрат(11) = 14,3389 [0,2148]

Модель 2: Пробит, использованы наблюдения 11-41 (Т = 31)
 Зависимая переменная: PresUS_1_0

	Коэффициент	Ст. ошибка	z	P-значение
const	6,66828	20,8901	0,3192	0,74957
Year	-0,00277574	0,0102903	-0,2697	0,78736
PresUS_1_0_1	0,171905	0,713399	0,2410	0,80958
PresUS_1_0_2	-0,793315	0,690227	-1,1494	0,25041
PresUS_1_0_3	-0,887974	0,667491	-1,3303	0,18341
PresUS_1_0_4	0,203741	0,674997	0,3018	0,76277
PresUS_1_0_5	-0,107105	0,672312	-0,1593	0,87343
PresUS_1_0_6	-0,774124	0,678556	-1,1408	0,25394
PresUS_1_0_7	0,619123	0,694201	0,8919	0,37247
PresUS_1_0_8	0,79107	0,671239	1,1785	0,23859
PresUS_1_0_9	-0,622303	0,673792	-0,9236	0,35570
PresUS_1_0_10	-0,764684	0,646623	-1,1826	0,23698
Среднее зав. перемен	0,516129	Ст. откл. зав. перемен		0,398309
R-квадрат Макфаддена	0,328786	Испр. R-квадрат		-0,230096
Лог. правдоподобие	-14,41192	Крит. Акаике		52,82384
Крит. Шварца	70,03168	Крит. Хеннана-Куинна		58,43316

Количество 'корректно предсказанных' случаев = 23 (74,2%)
 $f(\beta'x)$ для среднего значения независимых переменных = 0,398
 Критерий отношения правдоподобия: Хи-квадрат(11) = 14,119 [0,2265]

Модель 3: Тобит, использованы наблюдения 11-41 (Т = 31)
 Зависимая переменная: PresUS_1_0

	Коэффициент	Ст. ошибка	z	P-значение
const	7,96031	18,1226	0,4392	0,66048
Year	-0,00372172	0,00888998	-0,4186	0,67548
PresUS_1_0_1	0,136921	0,505278	0,2710	0,78641
PresUS_1_0_2	-0,350711	0,454862	-0,7710	0,44069
PresUS_1_0_3	-0,426251	0,523056	-0,8149	0,41512
PresUS_1_0_4	0,149234	0,465353	0,3207	0,74845
PresUS_1_0_5	0,0486054	0,436418	0,1114	0,91132
PresUS_1_0_6	-0,329703	0,461084	-0,7151	0,47457
PresUS_1_0_7	0,25159	0,485403	0,5183	0,60424
PresUS_1_0_8	0,433659	0,53488	0,8108	0,41750
PresUS_1_0_9	-0,412052	0,433133	-0,9513	0,34144
PresUS_1_0_10	-0,400537	0,491064	-0,8157	0,41470
Среднее зав. перемен	0,516129	Ст. откл. зав. перемен		0,508001
Цензур. наблюдения	15	Ст. ошибка модели		0,667506
Лог. правдоподобие	-25,21877	Крит. Акаике		76,43754
Крит. Шварца	95,07937	Крит. Хеннана-Куинна		82,51431
Параметр rho	0,106194	Стат. Дарбина-Вотсона		1,581477

Тест на нормальное распределение ошибок -
 Нулевая гипотеза: ошибки распределены по нормальному закону
 Тестовая статистика: Хи-квадрат(2) = 19,094
 р-значение = 7,14147e-005

Авторегрессия.

Модель 1: AR, использованы наблюдения 21-41 (Т = 21)
 Зависимая переменная: PresUS_1_0

	Коэффициент	Ст. ошибка	t- статистика	P-значение	
const	-6,5563	0,379436	-17,2791	<0,00001	***
Year	0,00355149	0,000188109	18,8799	<0,00001	***
PresUS_1_0_1	0,237723	0,0270115	8,8008	0,00001	***
PresUS_1_0_2	-0,283145	0,02036	-13,9069	<0,00001	***
PresUS_1_0_3	-0,107171	0,0130151	-8,2343	0,00002	***
PresUS_1_0_4	0,194313	0,0139871	13,8923	<0,00001	***
PresUS_1_0_5	-0,0931166	0,0165384	-5,6303	0,00032	***
PresUS_1_0_6	-0,21155	0,012458	-16,9811	<0,00001	***
PresUS_1_0_7	0,397638	0,0163263	24,3557	<0,00001	***
PresUS_1_0_8	0,436511	0,0174575	25,0042	<0,00001	***
PresUS_1_0_9	-0,324191	0,0184679	-17,5543	<0,00001	***
PresUS_1_0_10	-0,0495433	0,0243834	-2,0318	0,07272	*
u(-1)	-0,814596	0,0341201	-23,8744	<0,00001	***
u(-2)	-0,46557	0,0346145	-13,4501	<0,00001	***
u(-3)	-0,0172353	0,0218704	-0,7881	0,44731	
u(-4)	0,0497351	0,021443	2,3194	0,04062	**
u(-5)	-0,37699	0,0202037	-18,6595	<0,00001	***
u(-6)	-0,0677727	0,0220962	-3,0672	0,01071	**
u(-7)	0,0453258	0,0211263	2,1455	0,05508	*
u(-8)	-0,870828	0,0242765	-35,8713	<0,00001	***
u(-9)	-0,210293	0,0319479	-6,5824	0,00004	***
u(-10)	-0,59157	0,032151	-18,3997	<0,00001	***

Сумма AR-коэффициентов = -3,31979

Статистика, основанная на последней итерации вычисления параметра rho:			
Среднее зав. перемен	2,113136	Ст. откл. зав. перемен	0,728020
Сумма кв. остатков	0,012356	Ст. ошибка модели	0,037053
R-квадрат	0,997620	Испр. R-квадрат	0,994712
F(11, 9)	701,0876	P-значение (F)	8,31e-12
Параметр rho	0,454369	Стат. Дарбина-Вотсона	0,708093

Модель 2: Кохрана-Оркатта (Cochrane-Orcutt), использованы наблюдения 12-41 (Т = 30)

Зависимая переменная: PresUS_1_0

	Коэффициент	Ст. ошибка	t-статистика	P-значение
const	6,21834	19,976	0,3113	0,75916
Year	-0,00289568	0,00687564	-0,4212	0,67863
PresUS_1_0_1	0,572608	3,27975	0,1746	0,86335
PresUS_1_0_2	-0,445121	3,11006	-0,1431	0,88778
PresUS_1_0_3	0,126757	2,59098	0,0489	0,96152
PresUS_1_0_4	0,097217	1,31519	0,0739	0,94189
PresUS_1_0_5	0,0414935	0,232823	0,1782	0,86054
PresUS_1_0_6	-0,125823	0,268448	-0,4687	0,64491
PresUS_1_0_7	0,193014	0,521688	0,3700	0,71571
PresUS_1_0_8	0,111623	0,856893	0,1303	0,89780
PresUS_1_0_9	-0,370259	0,248299	-1,4912	0,15323
PresUS_1_0_10	-0,205227	1,37077	-0,1497	0,88265

Статистика, основанная на последней итерации вычисления параметра rho:

Среднее зав. перемен	0,533333	Ст. откл. зав. перемен	0,507416
Сумма кв. остатков	3,240524	Ст. ошибка модели	0,424298
R-квадрат	0,566001	Испр. R-квадрат	0,300780
F(11, 18)	3,150250	P-значение (F)	0,015068
Параметр rho	0,035230	Стат. Дарбина-Вотсона	1,922085

Исключая константу, наибольшее р-значение получено для переменной 7 PresUS_1_0_3)

Модель 3: VAR система, порядок лага 2

Метод оценки - МНК, наблюдения 3-41 (Т = 39)

Лог. правдоподобие = -25,707673

Определитель ковариационной матрицы = 0,21881364

Крит. Акаике = 1,4722

Крит. Шварца = 1,6002

Крит. Хеннана-Куинна = 1,5181

Портмане-тест (Portmanteau): LB(9) = 5,0747, Ст. свободы = 7 [0,6508]

Уравнение 1: PresUS_1_0

Робастные оценки стандартных ошибок (с поправкой на гетероскедастичность), вариант HC1

	Коэффициент	Ст. ошибка	t-статистика	P-значение
PresUS_1_0_1	0,209885	0,157734	1,3306	0,19168
PresUS_1_0_2	-0,293893	0,150327	-1,9550	0,05838 *
Year	0,00032852	7,50591e-05	4,3768	0,00010 ***
Среднее зав. перемен	0,589744	Ст. откл. зав. перемен		0,498310
Сумма кв. остатков	8,533732	Ст. ошибка модели		0,486876
R-квадрат	0,628968	Испр. R-квадрат		0,608355
F(3, 36)	20,34224	P-значение (F)		6,99e-08
Параметр rho	-0,002520	Стат. Дарбина-Вотсона		1,971441

F-тесты для нулевых ограничений:

Все лаги для PresUS_1_0 $F(2, 36) = 2,813 [0,0733]$

Все переменные, лаг 2 $F(1, 36) = 3,8221 [0,0584]$

Для всей системы

Нулевая гипотеза: самый длинный лаг 1

Альтернативная гипотеза: самый длинный лаг 2

Критерий отношения правдоподобия: Хи-квадрат(1) = 3,48597 [0,0619]

Сравнение информационных критериев:

Лаг порядка 2: AIC = 1,47219, BIC = 1,60015, HQC = 1,51810

Лаг порядка 1: AIC = 1,51029, BIC = 1,59560, HQC = 1,54090