

ИСПОЛЬЗОВАНИЕ СУПЕРКОМПЬЮТЕРА
ДЛЯ ЭКСПЕРИМЕНТАЛЬНОГО
ИССЛЕДОВАНИЯ МЕТОДА
ПРОГНОЗИРОВАНИЯ, ОСНОВАННОГО
НА УНИВЕРСАЛЬНЫХ КОДАХ

Приставка Павел,
аспирант СибГУТИ

Временные ряды и прогнозирование в жизни

- ◎ Экономика
 - Курсы валют
 - Индексы
 - Котировки акций
- ◎ Физические явления
 - Излучение
 - Геомагнитная активность
 - Метеорология
- ◎ Спортивная статистика
- ◎ Общество
 - Демография
 - Авиаперевозки
 - Уровень преступности

Понятия и определения

Код U называется универсальным, если для любого стационарного и эргодического источника верно:

$$\lim_{t \rightarrow \infty} |U(x_1 \dots x_t)| / t = H(P)$$

Мера R :

$$R(x_1 \dots x_t) = \sum_{i=0}^{\infty} \omega_{i+1} K_i(x_1 \dots x_t)$$

Код Б.Я. Рябко:

$$\tilde{R}(x_1 \dots x_t) = \log \left(\sum_{i=0}^{\infty} \omega_{i+1} K_i(x_1 \dots x_t) \right)$$

Прогнозирование

Источник порождает значения из конечного алфавита

Для любого элемента a из алфавита источника:

$$R(a | x_1 \dots x_t) = \frac{R(x_1 \dots x_t a)}{R(x_1 \dots x_t)}$$

Прогнозирование

Источник принимает значения из непрерывного интервала

Пусть X_t - стационарный и эргодический процесс, принимающий значения из интервала Ω и $\{\Pi_n\}$ – возрастающая последовательность конечных разбиений Ω . Пусть совместное распределение вероятности $P_n(X_1, \dots, X_n)$ имеет плотность $p_n(X_1, \dots, X_n)$ по отношению к сигма-конечной мере M_n . Для целых s и n определим соответствующую плотность как:

$$p^s(x_1, \dots, x_n) = P(x_1^{[s]}, \dots, x_n^{[s]}) / M_n(x_1^{[s]}, \dots, x_n^{[s]})$$

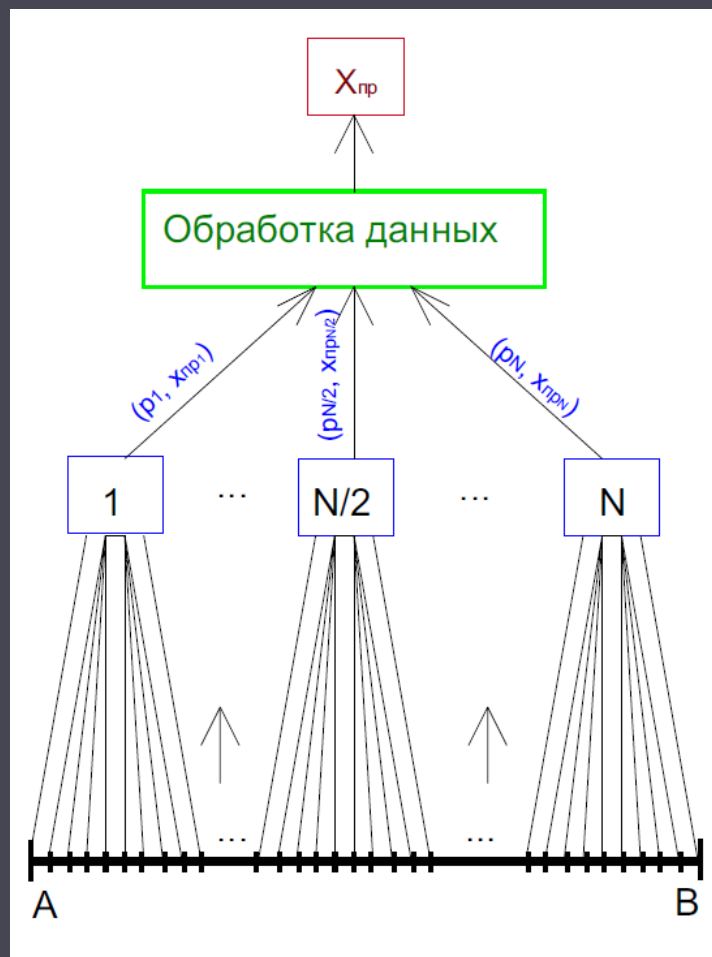
Оценка неизвестной плотности:

$$r_U(x_1 \dots x_t) = \sum_{i=1}^{\infty} \omega_i R(x_1^{[i]} \dots x_t^{[i]}) / M_t(x_1^{[i]} \dots x_t^{[i]})$$

Тогда, оценка условной вероятности:

$$r_U(a | x_1 \dots x_m) = \frac{r_U(x_1 \dots x_m a)}{r_U(x_1 \dots x_m)}$$

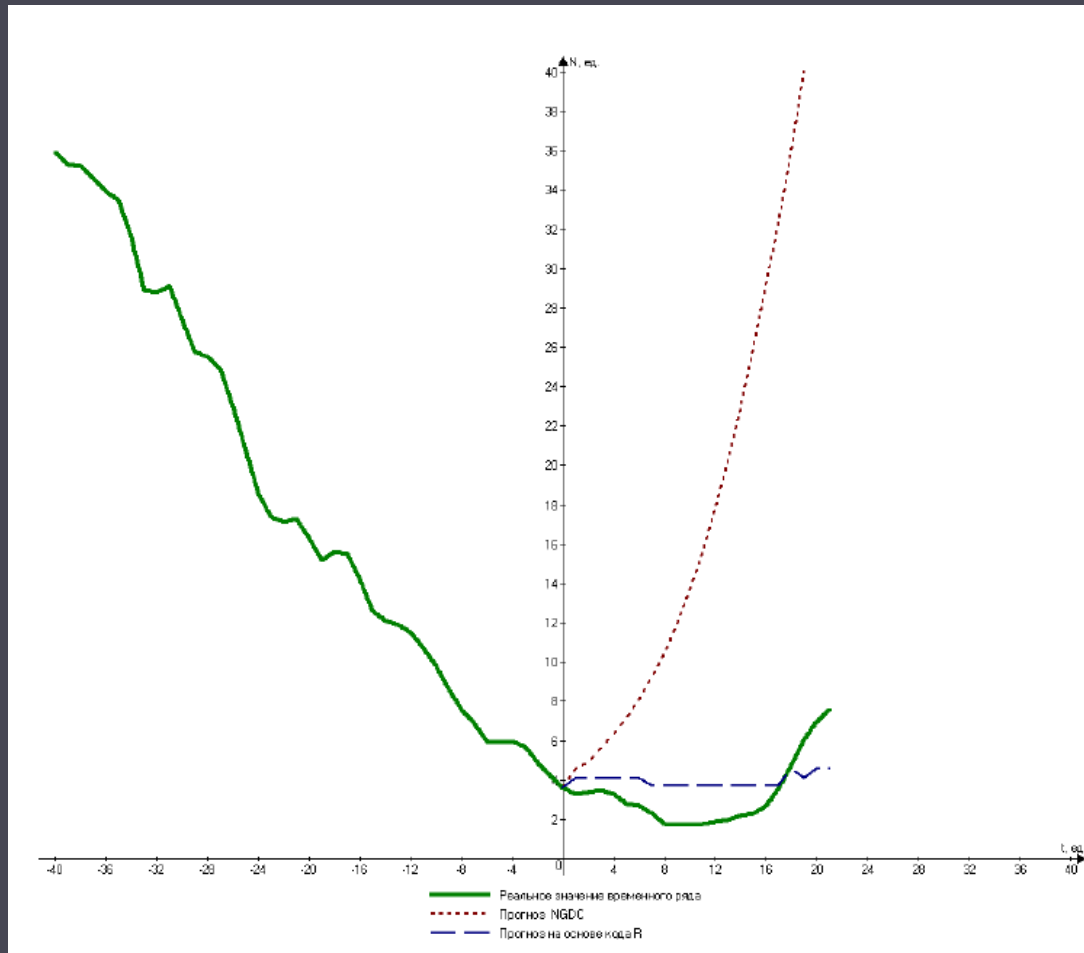
Распараллеливание вычислений



Экспериментальные результаты

Временной ряд	Диапазон значений процесса	Длина						
		500	700	1000	1200	2000	3000	4000
Среднемесячное число солнечных пятен	[0; 256]	6.54	2.56	9.58	15.85	21.7	19.63	н/д
Сглаженное среднемесячное число солнечных пятен	[0; 210]	1.5	1.1	1.99	0.77	3.36	2.56	н/д
Абсолютное ежедневное солнечное излучение	[50; 300]	1.17	1.17	2.71	5.52	8.35	1.72	1.45
Абсолютное ежемесячное солнечное излучение	[580; 2540]	211.29	45.88	н/д	н/д	н/д	н/д	н/д
DST индекс	[-589;44]	3.08	2.4	1.88	4.18	3.68	3.24	3.6

Экспериментальные результаты



R	Длина		
	500	1000	2000
NGDC	1.75	1.87	2.23
NGDC	16.39		

Экспериментальные результаты

Длина временно- го ряда	Диапазон значений процесса	Statistica 8.0	R
500	[-63;24]	4.36	3.08
1000	[-111;24]	7.56	1.88
2000	[-129;28]	6.16	3.68
3000	[-129;44]	10	3.24
4000	[-589;24]	18.24	3.6

