

## ВИКОРИСТАННЯ АНАЛІТИЧНИХ ТЕХНОЛОГІЙ ДЛЯ ПРОГНОЗУВАННЯ ДИНАМІЧНИХ ПРОЦЕСІВ У ФІНАНСОВІЙ СФЕРІ

Інститут комп'ютерних технологій  
Національного авіаційного університету

*Розглянуто питання представлення часових дискретних даних та узагальнення їх з метою спрощення обробки. Визначено результати використання методу поліноміальної екстраполяції для прогнозування даних рядів та дано оцінки достовірності прогнозів. Створений клас даних в середовищі Microsoft Visual Studio для розв'язання систем будь-якого порядку*

### Вступ

Аналітичні технології – це методики, які на основі деяких моделей, алгоритмів, математичних теорем дозволяють за відомими даними оцінити значення невідомих характеристик та параметрів.

Найпростішим прикладом є теорема Піфагора, яка дозволяє за довжинами сторін прямокутника визначити довжину діагоналі. Іншим прикладом аналітичної технології є способи, за допомогою яких оброблює інформацію людський мозок. Навіть мозок дитини здатний вирішувати такі задачі, які не під силу сучасним комп'ютерам, як, наприклад, розпізнавання обличчя в натовпі чи контроль десятків м'язів при грі у футбол. Унікальність мозку складається в тому, що він здатний вчитися розв'язанню нових задач (гри в шамати, водінню автомобіля), але погано пристосований для розв'язання великих об'ємів інформації (обрахування кореня з числа 463761). На практиці часто зустрічаються більш складні задачі ніж вирахування квадратного кореня. Таким чином людині для розв'язання цих задача потрібні додаткові методики та інструменти.

Аналітичні технології в першу чергу потрібні людям, які приймають важливі рішення – керівники, аналітики, експерти, консультанти. Прибуток компанії в більшості випадків залежить від якості цих рішень – точності прогнозів, оптимально обраних стратегій [1].

### Підготовка даних до аналізу

Для узагальнення даних часових рядів використовується інструмент «бар» («японська свічка»). Кожний бар відповідає якомусь інтервалу часу (30 хвилин, година, чотири години, день). Бар складається з 6 елементів:

1. Період бара ( $t$ ).
2. Відкриття (*open*).
3. Закриття (*close*).
4. Максимальне значення (*high*).
5. Мінімальне значення (*low*).
6. Напрямок динаміки.

### Формула прогнозу

Для прогнозування ми використовуємо деякі принципи аналітичної геометрії

Точки на площині можна описати за допомогою функції  $n$ -ого степеня. Для цього необхідно мати  $n+1$  точок. 2 точки визначають пряму, 3 – параболу, 4 – кубічну параболу і т.д.

Покажемо цей принцип на прикладі параболи (кривої 2-го порядку):

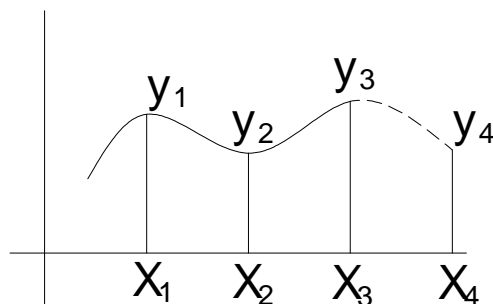


Рис. 1.

Є деяка функція  $y(x)$ . Існує три значення  $X$  і відповідні їм значення  $Y$ . Задамо функції  $y(x)$  в загальному вигляді:

$$y(x) = a_0 \cdot x^2 + a_1 \cdot x + a_3$$

Підставивши значення  $X$  ми отримаємо систему рівнянь:

$$\begin{cases} a_0 \cdot x_1^2 + a_1 \cdot x_1 + a_3 = y_1 \\ a_0 \cdot x_2^2 + a_1 \cdot x_2 + a_3 = y_2 \\ a_0 \cdot x_3^2 + a_1 \cdot x_3 + a_3 = y_1 \end{cases}$$

Розв'язком системи рівнянь є значення коефіцієнтів  $a$ , які визначають нашу функції  $y(x)$ .

Після цього ми можемо отримати значення  $y(x_4)$ .

Таким чином ми можемо виконувати прогнозування 4 показника при наявності трьох попередніх. У такому випадку ві-

домі три показники – це значення  $y$ , а в якості значення  $X$  – числа, що відповідають моментам часу. Наприклад: 1, 2, 3 (якщо показники були отримані через однакові елементи часу).

Для автоматизації процесу розрахунку наступного значення ми використовуємо середовище програмування *Microsoft Visual Studio 2005* і мову програмування *C#*. Для полегшення ми створили клас, що дозволяє задати матрицю коефіцієнтів і розв'язати систему рівнянь матричним методом будь-якої (!) складності. Це дозволить в майбутньому використовувати прогнози з кривими різного порядку.

### Спосіб ведення статистики

Для тестової моделі використаємо дані за 2008 рік. Ми беремо добові свічки. Прогноз справдився, якщо прогнозна точка попала в тіло свічки.

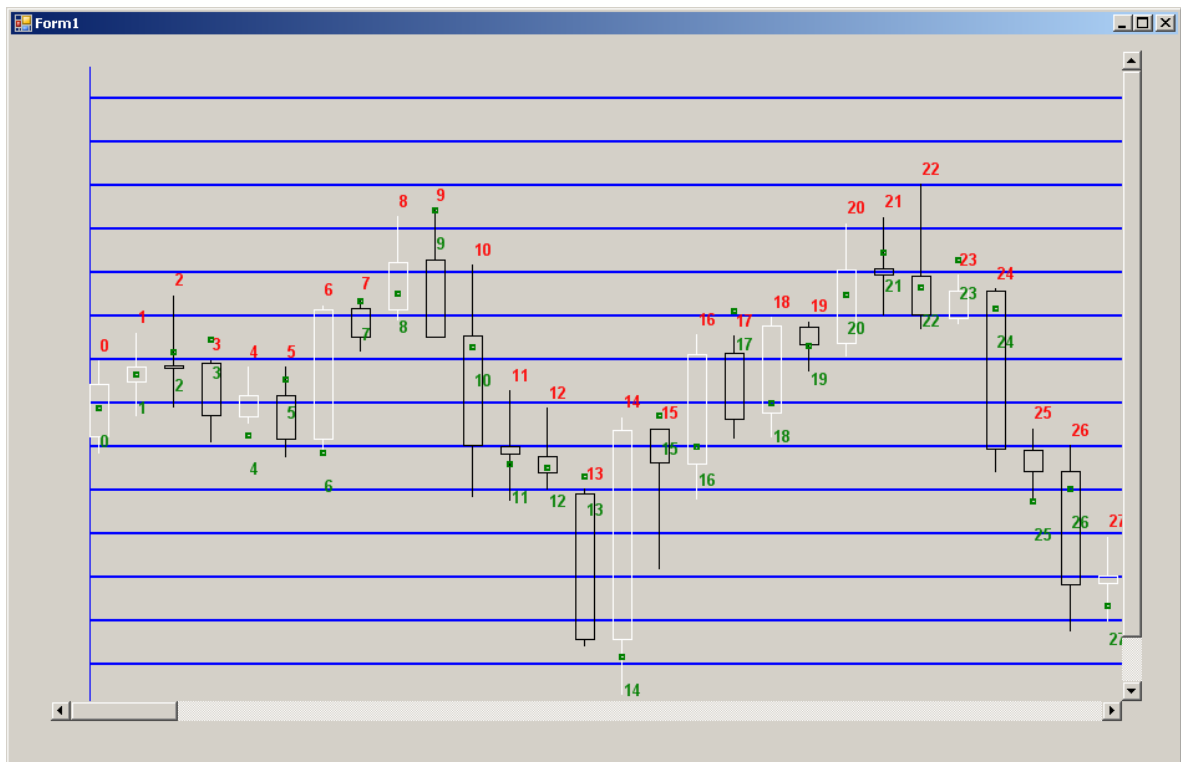


Рис. 2. Скріншот результатів прогнозу

На даному рисунку присутня ілюстрація свічок.

Над кожною свічкою встановлений номер для відслідковування правильності роботи.

Для кожної свічки зроблений прогноз, нижче якого стоїть його номер.

Слід зазначити, що якості значень вибрано середину між максимумом і мінімумом.

Для перших  $n+1$  свічок прогноз не робиться і встановлюється рівним середині свічки. ( $n$  – порядок кривої).

Прогноз вважається вдалим, якщо він попадає в тіло свічки (прогнози 5, 7, 8, 10) і невдалим, якщо не попав (прогнози 3, 4, 6, 13) [2].

Спочатку виведемо статистику по вдалості в залежності від таких параметрів:

1. Порядок кривої (від 2 до 7).
2. Початкові значення (*open*, *close*, *high*, *low*,  $(high+low)/2$ ).

Слід зазначити, що кількість свічок в кожному році – 260. (кількість робочих днів)

Таблиця 1. Статистика фіксування прогнозу

Порядок кривої \ Роки	2	3	4	5	6	7
2008	110	67	40	21	3	0
2007	99	64	38	17	0	0
2006	94	55	26	20	0	0
2005	92	57	35	18	0	0
2004	84	54	35	16	0	0

В даному досліді використовувалися початкові значення рівні  $(high+low)/2$ .

З цього досліді ми можемо зробити висновок про зниження ефективності з ростом порядку кривої, тому надалі будемо використовувати лише 2 та 3 порядок.

Таблиця 2. Статистика фіксування прогнозу по різним вхідним даним протягом 2004 р.

Порядок кривої \ Роки	2	3	4	5	6
2004	84	54	35	16	0
2004 – закр	72	35	23	15	1
2004 – хай	60	47	22	9	0
2004 – лоу	71	40	29	13	1
2004 – опен	52	36	21	9	0
2004 – $(open+close)/2$	104	81	55	38	0

Звідси видно, що найбільшій вірогідності ми досягаємо, коли у якості вхід-

них даних використовуємо значення середини тіла  $((відкриття+закриття)/2)$ .

Для більшої точності проведемо дослідження на даних іншого року. Наприклад 2005.

Таблиця 3. Статистика фіксування прогнозу по різним вхідним даним протягом 2005 р.

Порядок кривої \ Роки	2	3	4	5	6
2005	92	57	35	18	0
2005 – закр	65	33	19	10	0
2005 – хай	64	39	23	10	1
2005 – лоу	94	52	29	18	0
2005 – опен	54	30	20	6	0
2005 – $(open+close)/2$	135	92	72	35	0

Таким чином ми визначили найбільш оптимальні вхідні параметри. Це:

1. 2 та 3 порядок кривої.
2. Значення між відкриття та закриттям свічки.

### Величина прогнозу

Так економічні дані не є абстрактними, їх значення коливаються в якомусь досить широкому інтервалі, але не виходять за рамки границь і не мають значних скачків.

Наприклад, ціна не може підскочити в один день більше ніж на 20%. Можливий стрімкий рух угору (або вниз), але є обмеження на швидкість цього руху.

Оскільки процес зміни досить динамічний і відчуває вплив високої кількості учасників, то завжди існують невеликі коливання, які не змінюють загальну тенденцію, тому варто прийняти за увагу лише ті прогнози, які дають значну зміну і відкинути прогнози, які прогнозують незначну зміну показника.

Для визначення цих значень проведемо дослідження. Виведемо статистичний розподіл по довжині свічок (*high-low*).

Найбільша кількість свічок по довжині розташовується в інтервалі від 0,0030-0,0040.

Тому прийемо стартову мінімальну силу прогнозу – 0,0020.

Тобто в розрахунок будемо брати тільки прогнози, які дають рух більше ніж 0,0020 [3].

Таблиця 4. Кількість довжин добових свічок за 2006 р.

Діапазон довжини свічки, від	Кількість
0,0000	0
0,0010	70
0,0020	158
0,0030	220
0,0040	175
0,0050	162
0,0060	154
0,0070	96
0,0080	105
0,0090	99
0,0100	70
0,0110	59
0,0120	34
0,0130	30
0,0140	23
0,0150	24
0,0160	11
0,0170	13
0,0180	12
0,0190	7
0,0200	4
0,0210	3
0,0220	4
Більше 0,0230	18

### Тенденція та її зміна

Враховуючи хвильовий рух показників, тобто існують хвилі росту або спаду довжину в декілька свічок підряд є смисл обчислити показник даної тенденції. Одним з методів визначення є середня. Середня обчислюється як середнє арифметичне декількох попередніх свічок. В залежності від кількості цих свічок визначається довжина і тип середньої. Довжина – кількість свічок. Якщо довжина від 3 до 6 – це коротка середня, якщо від 15 до 20 – довга. Для кожного окремого показника підбирається своє значення довжина, але принцип один: довга в 4-5 разів довша короткої.

Дані інструмент (середні) будуть показувати тенденцію зміни показника. Покази будуть іти з запізненням. Для довгої середньої запізнення буде більшим,

тому коли маємо обидва інструменти, то при зміні короткої середньої потрібно чекати зміну довгої. Даний висновок дає нам можливість оцінювати тенденцію та зміну її напрямку.

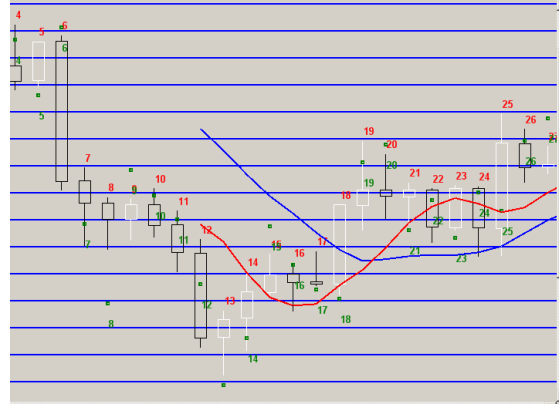


Рис. 3. Довга (синя) і коротка (червона) середні. Довжина 4 та 12

### Висновки

1. Виконувати поліноміальну екстраполяція раціонально за допомогою кривих 2 та 3 порядку. Це зменшить кількість обчислень та збільшить продуктивність системи.

2. У якості вхідних даних важливо вибрати правильні точки. У даному випадку середина тіла свічки.

3. Потрібно враховувати величину зміни за прогнозам. Дуже великі або дуже малі прогнози не вірогідними і підлягають виключенню.

### Список літератури

1. Джон Э. Ханк, Дин У. Уичерн, Артур Дж. Райтс. Бизнес-прогнозирование. Вільямс, 2006. – 663 с.
2. Эндрю Сигел. Практическая бизнес-статистика. Вільямс, 2008. – 556 с.
3. Хемди А. Практическая бизнес-статистика. Вільямс, 2005. – 912 с.