УДК 519.2

Регрессионный анализ характера зависимости толщины пленки углерода от температуры: тезисы доклада

А. М. Ревякин, А. В. Бардушкин

Национальный исследовательский университет «МИЭТ», Москва, Россия

i170k@yandex.ru

Regression Analysis of Type of the Carbon Film Thickness Dependence on Temperature

A. M. Revyakin, A. V. Bardushkin

National Research University of Electronic Technology, Moscow, Russia

i170k@yandex.ru

The authors did build regression model of the carbon film thickness change depending on temperature. They have shown a strong linear relation of the studied factors. The authors did obtain interval estimates for the main regression parameters and established the significance of their model. They did make a conclusion about possibility to predict the characteristics of graphite nanoscale films based on the regression equations.

Keywords: regression model; hypothesis testing; model significance.

Исследование уникальных свойств графена является сегодня одним из наиболее перспективных направлений разработки микро- и наноэлектронных устройств. Графен — это двумерная структура, в которой атомы углерода выстроены в форме правильных шестиугольников. Графен обладает рядом свойств, в том числе является полупроводником. Однако электронные свойства графена быстро меняются при увеличении количества слоев.

Структуры с числом слоев больше десяти называют пленками графита. Воспользуемся данными исследований наноструктур на основе слоев графена [1]. При разработке технологии формирования таких наноструктур синтез наноразмерных пленок графита осуществлялся на поверхности никелевого катализатора на кремниевой подложке.

Известны размеры Y (нм) толщины пленки углерода при изменении температуры X (°С) (см. таблицу). Требуется определить тип зависимости, степень коррелированности компонент двумерного вектора $(X; Y)$, показать статистическую значимость построенной модели.

X	600	610	620	630	640	650	660	670	680	690	700
Y	0,852	0,828	0,801	0,770	0,736	0,697	0,654	0,606	0,552	0,493	0,428

Объем двумерной выборки n = 11. Диаграмма рассеивания представлена на рисунке.

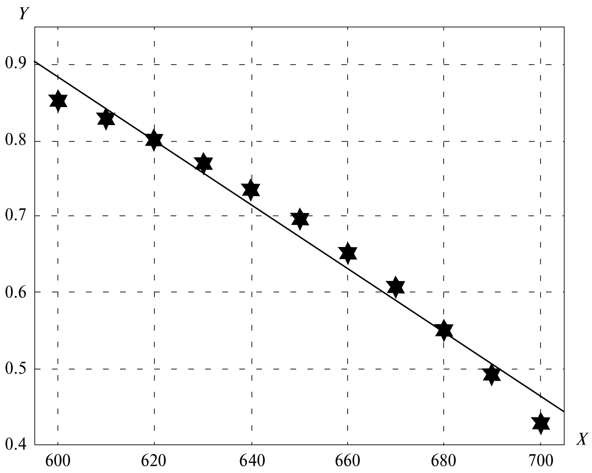

Диаграмма рассеивания и график уравнения выборочной линейной регрессии

Проведем вычисление параметров выборки для негруппированных данных [2; 3; 4]: выборочные средние $\overline{x} = 650$, $\overline{y} = 0{,}6743$, несмещенные оценки дисперсий $s_X^2 = 1100$, $s_Y^2 = 0{,}0198$, выборочная корреляция $\tilde{\rho}_{X,Y} = -0{,}9886$.

Проверим гипотезу о наличии корреляции. Для небольшой по объему выборки выберем статистику

$$U = \frac{\operatorname{arth}\tilde{\rho}_{X,Y} - \operatorname{arth}\rho_0}{1/\sqrt{n-3}}$$

для проверки гипотезы $H_0: \rho_{X,Y} = \rho_0$, где

$$\operatorname{arth} x = \frac{1}{2}\ln\frac{1+x}{1-x},$$

а $\rho_0 = 0$, против любой из альтернатив. Выборочное значение статистики $u_в = -7{,}2937$, квантиль $u_{0,975} = 1{,}96$. Поскольку $|u_в| > u_{0,975}$, то гипотеза H_0 отклоняется в пользу альтернативной гипотезы. Корреляция значима на уровне $\alpha = 0{,}05$.

При малом объеме выборки для коэффициента корреляции необходима еще интервальная оценка. Соответствующий доверительный интервал:

$$-0{,}9968 < \rho < -0{,}9504.$$

Для негруппированной двумерной выборки составим уравнения прямых регрессий Y на x:

$$y(x) = \overline{y} + \tilde{\rho}_{X,Y}\frac{s_Y}{s_X}(x - \overline{x})$$

и X на y:

$$x(y) = \overline{x} + \tilde{\rho}_{X,Y}\frac{s_X}{s_Y}(y - \overline{y}).$$

Получим $y = 3{,}4025 - 0{,}0042x$, $x = 806{,}9885 - 232{,}8265y$. Прямые $y = \tilde{a}x + \tilde{b}$ и $x = \tilde{a}'y + \tilde{b}'$ пересекаются в точке с координатами (650;0,06743), причем угол между ними практически нулевой, так как коэффициент корреляции близок к минус единице. На рисунке показан график прямой регрессии Y на x.

Качество аппроксимации результатов наблюдений выборочной регрессией $y = \tilde{a}x + \tilde{b}$ определяется величиной остаточной дисперсии, вычисляемой по формуле

$$s^2 = \frac{Q_e}{n-2}.$$

Остаточная сумма квадратов $Q_e = \sum_{i=1}^{n}(y_i - \tilde{y}_i)^2$ равна сумме квадратов разностей между наблюдаемыми значениями переменной Y при $x = x_i$ и расчетными значениями $\tilde{y}_i = \tilde{a}x_i + \tilde{b}$. Эти разности называются остатками и в данном случае имеют значения: –0,0142; 0,0008; 0,0118; 0,0198; 0,0227; 0,0217; 0,0157; 0,0036; –0,0364. Коэффициент детерминации характеризует ту долю разброса результатов наблюдений относительно горизонтальной прямой $y = \overline{y}$, которая объясняется выборочной регрессией $y = \tilde{a}x + \tilde{b}$. Вычисленное нами значение коэффициента детерминации $R^2 = 0{,}9772$ показывает, что уравнение регрессии $y = 3{,}4025 - 0{,}0042x$ на 97,72 % объясняет общий разброс результатов наблюдений относительно горизонтальной прямой $y = 0{,}6743$.

Проверим значимость линейной регрессии Y на x при $\alpha = 0{,}05$. Построим доверительный интервал для коэффициента $\widetilde{a}$ [3]. Гипотеза H_0: $a = 0$ отклоняется на уровне значимости $\alpha = 0{,}05$, так как доверительный интервал

$$-0{,}0047 < a < -0{,}0037$$

не накрывает нуль с доверительной вероятностью 0,95.

Этот же результат можно получить, используя для проверки гипотезы H_0: $a = 0$ статистику $F = (n - 2) \cdot Q_R/Q_e$, где $Q_R = \sum\limits_{i=1}^{n} (\widetilde{y}_i - \bar{y})^2$ — сумма квадратов, обусловленная регрессией. Выборочное значение статистики $f_в = 386{,}3656$. Если

$$F_{\frac{\alpha}{2}}(1; n-2) < f_в < F_{1-\frac{\alpha}{2}}(1; n-2),$$

то нет оснований отвергать гипотезу H_0: $a = 0$; иначе H_0 отклоняется (здесь $F_p(1; n - 2)$ — квантиль распределения Фишера с 1 и $n - 2$ степенями свободы). Поскольку $f_в \notin (0{,}0010; 7{,}5709)$, то гипотеза H_0: $a = 0$ отклоняется на уровне значимости $\alpha = 0{,}05$. Таким образом, линейная регрессия Y на x статистически значима.

При попытке построить полиномиальную аппроксимацию данных коэффициент при старшей степени для квадратичной функции оказался практически нулевым.

Таким образом, наблюдаемая зависимость (см. таблицу) хорошо описывается линейной моделью. Полученные нами уравнения регрессий могут использоваться для прогнозирования толщины пленки графита при небольших изменениях температуры, при условии сохранения химических и физических состояний материалов.

Графитовые пленки, наряду с электронными, обладают ценными теплофизическими и механическими свойствами (в том числе упругостью), поэтому их изучение актуально в первую очередь при разработке оптоэлектронных изделий и в целом для наноэлектроники.

Литература

1. ***Левин Д. Д.*** Разработка и исследование технологии формирования наноструктур с проводящим каналом на основе слоев графена: автореф. дис. ... канд. техн. наук. М., 2015. 21 с.

2. ***Чистяков В. П.*** Курс теории вероятностей. 8-е изд., испр. М.: Ленанд, 2015. 304 с.

3. Лабораторный практикум по курсу «Теория вероятностей и математическая статистика» / В. В. Бардушкин, В. В. Лесин, В. Н. Земсков, Н. Н. Мустафин. М.: МИЭТ, 2009. 116 с.

4. Задания для выполнения лабораторных и индивидуальных работ по курсу «Теория вероятностей и математическая статистика» с использованием пакета MATLAB / В. В. Бардушкин, И. В. Бардушкина, В. В. Лесин, А. М. Ревякин // Проектирование инженерных и научных приложений в среде MATLAB: Мат-лы V Междунар. науч. конф. (г. Харьков, 11—13 мая 2011 г.) / Сост. В. В. Замаруев. Харьков: БЭТ, 2011. С. 471—533.

Ревякин Александр Михайлович — кандидат физико-математических наук, доцент кафедры высшей математики № 2 Национального исследовательского университета «МИЭТ» (Россия, 124498, Москва, г. Зеленоград, пл. Шокина, д. 1), *arevyakin@mail.ru*

Бардушкин Андрей Владимирович — студент группы ЭКТ-43 Национального исследовательского университета «МИЭТ» (Россия, 124498, Москва, г. Зеленоград, пл. Шокина, д. 1), *i170k@yandex.ru*

References

1. Levin D. D. Razrabotka i issledovanie tekhnologii formirovaniya nanostruktur s provodyashchim kanalom na osnove sloev grafena, avtoref. dis. ... kand. tekhn. nauk (Development and Study of Technology of Nanostructures Formation with Conducting Channel Based on Graphene Layers, Extended Abstract of Cand. Sci. (Engineering) Dissertation), M., 2015, 21 p.

2. Chistyakov V. P. Kurs teorii veroyatnostei (Course of Theory of Probability), 8-e izd., ispr., M., Lenand, 2015, 304 p.

3. Laboratornyi praktikum po kursu "Teoriya veroyatnostei i matematicheskaya statistika" (Laboratory Practicum on Theory of Probability and Mathematical Statistics), by V. V. Bardushkin, V. V. Lesin, V. N. Zemskov, N. N. Mustafin, M., MIET, 2009, 116 p.

4. Zadaniya dlya vypolneniya laboratornykh i individual'nykh rabot po kursu "Teoriya veroyatnostei i matematicheskaya statistika" s ispol'zovaniem paketa MATLAB (Tasks for Laboratory and Independent Work in "Theory of Probability and Mathematical Statistics" Course Using MATLAB Software), by V. V. Bardushkin, I. V. Bardushkina, V. V. Lesin, A. M. Revyakin, *Proektirovanie inzhenernykh i nauchnykh prilozhenii v srede MATLAB: Mat-ly V Mezhdunar. nauch. konf. (g. Khar'kov, 11—13 maya 2011 g.),* Sost. V. V. Zamaruev, Khar'kov, BET, 2011, pp. 471—533.

Revyakin Alexander M., Ph.D. of physical and mathematical sciences, associate professor, associate professor of Higher Mathematics Department No. 2, National Research University of Electronic Technology (Shokin Square, 1, 124498, Moscow, Zelenograd, Russia), *arevyakin@mail.ru*

Bardushkin Andrey V., student of EKT-43 group, National Research University of Electronic Technology (Shokin Square, 1, 124498, Moscow, Zelenograd, Russia), *i170k@yandex.ru*