

ПРИКЛАДНАЯ МАТЕМАТИКА

УДК 519.25, 537.533.2
MSC 62J05

Метод случайного поиска при оценке параметров сигнала эмиссионной системы*

М. И. Варяонь, Е. М. Виноградова, А. Ю. Антонов

Санкт-Петербургский государственный университет, Российская Федерация,
199034, Санкт-Петербург, Университетская наб., 7–9

Для цитирования: *Варяонь М. И., Виноградова Е. М., Антонов А. Ю.* Метод случайного поиска при оценке параметров сигнала эмиссионной системы // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2021. Т. 17. Вып. 3. С. 228–239. <https://doi.org/10.21638/11701/spbu10.2021.301>

В рамках компьютерного статистического эксперимента описана тестовая задача идентификации параметров сигнала полевой электронной эмиссии с помощью регрессионной модели, основанной на законе Фаулера — Нордгейма. Были применены два подхода к определению оценок параметров — метод наименьших квадратов и случайный поиск с обучением. Показано, что погрешностью случайного поиска можно пренебречь, если при рассматриваемых соотношениях уровня шума к мощности сигнала количество статистических испытаний составит величину порядка 10^3 . Полученный результат позволяет расширить класс используемых для идентификации отклика функционалов, не меняя при этом метод. Отмечены преимущества метода случайного поиска для предложенной задачи и перспективы его применимости для задач в более общей постановке.

Ключевые слова: полевая электронная эмиссия, вольт-амперная характеристика, регрессионная модель, метод наименьших квадратов, метод случайного поиска.

1. Введение. Источники заряженных частиц, работающие в режиме полевой электронной эмиссии (ПЭЭ), давно используются в различных областях науки и техники. Это касается как вакуумных приборов и устройств [1–3], так и материаловедения в целом [4].

Запишем связь силы эмиссионного тока I с напряжением V между катодом и анодом в виде p -параметрической ($p = 2$) зависимости [5]

$$I(V; \mathbf{q}) = I_0 q_1 \left(\frac{V}{V_0} \right)^2 \exp \left[-q_2 \frac{V_0}{V} \right]. \quad (1)$$

* Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант № 20-07-01086).

© Санкт-Петербургский государственный университет, 2021

Это представление является следствием теории Фаулера — Нордгейма (ФН) [6]. Здесь компоненты вектора \mathbf{q} в нулевом приближении — постоянные величины, определяемые материалом катода и геометрическими параметрами системы. Калибровочные значения V_0 и I_0 призваны сделать характеристики q_1 и q_2 безразмерными. В ходе эксперимента \mathbf{q} оценивается по результатам наблюдений. Область значений параметров обозначим как $Q = \{(q_1, q_2) : q_1 > 0, q_2 > 0\}$. В данной работе не будем касаться имеющих в литературе отклонений от теории ФН и, как следствие, моделей с большим ($p > 2$) числом параметров [7, 8]. Тем не менее рассматриваемый подход может быть легко распространен на общий случай.

Осуществим следующее преобразование, которое называется *координатами ФН*:

$$X = \frac{V_0}{V}, \quad Y = \lg \left[\frac{I}{V^2} \frac{V_0^2}{I_0} \right], \quad \vartheta_1 = \lg q_1, \quad \vartheta_2 = -\frac{q_2}{\ln 10}. \quad (2)$$

Оно позволяет записать (1) в линейном виде как по параметрам $\vartheta = (\vartheta_1, \vartheta_2)^T \in \Theta$, так и по фактору X :

$$Y = \vartheta_1 + \vartheta_2 X, \quad (3)$$

где $\Theta = \{(\vartheta_1, \vartheta_2) : \vartheta_1 \in \mathbb{R}, \vartheta_2 < 0\}$ — отображение Q с помощью (2). Для систем одинаковой конфигурации наклон вольт-амперной характеристики (ВАХ) зависит от работы выхода электрона — крайне важной для эмиссионной электроники величины. Методика определения работы выхода по наклону ВАХ настолько отточена, что служит отправной точкой для дальнейших исследований в данной области [9].

Преобразование (2) дает простую регрессионную модель (3). Рассмотрим ее. Измерения силы тока, как и соответствующего напряжения, неизбежно содержат шум. Несложно показать, что ошибки измерения V можно до некоторой степени игнорировать, рассматривая их как составляющую аддитивной погрешности для силы тока: $\tilde{I} = I + I_0 \varepsilon$ [10]. Понятно, что переход к координатам ФН затрагивает и шум.

Применим метод наименьших квадратов (МНК) для поиска оптимальных оценок $\hat{\vartheta}$. В этом случае минимизируемый функционал имеет вид

$$J = J(\vartheta; \mathbf{X}, \tilde{\mathbf{Y}}) = \sum_{i=1}^N \left(\tilde{Y}_i - \vartheta_1 - \vartheta_2 X_i \right)^2, \quad (4)$$

а соответствующее решение представим следующим образом:

$$\hat{\vartheta} = \mathbf{M}^{-1} \mathbf{b}, \quad \mathbf{M} = \begin{pmatrix} N & \sum_{i=1}^N X_i \\ \sum_{i=1}^N X_i & \sum_{i=1}^N X_i^2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} \sum_{i=1}^N \tilde{Y}_i \\ \sum_{i=1}^N X_i \tilde{Y}_i \end{pmatrix},$$

$$\hat{q}_2 = 10^{\hat{\vartheta}_1}, \quad \hat{q}_2 = -\hat{\vartheta}_1 \ln 10.$$

Очевидно, что $\hat{\vartheta}$ естественным образом оказывается также случайной величиной. Оценивание влечет за собой появление понятия остатков регрессионной модели:

$$\hat{\varepsilon}_i = \tilde{Y}_i - \hat{\vartheta}_1 - \hat{\vartheta}_2 X_i. \quad (5)$$

Известно [11], что $\hat{\vartheta}$ будет оценкой максимального правдоподобия, если $\{\hat{\varepsilon}_i\}$ удовлетворяют следующим условиям. Так, остатки должны быть независимы и одинаково

нормально распределены. Если в выборке $\{\hat{\epsilon}_i\}$ наблюдается автокорреляция, гетероскедастичность или отклонение от нормальности в пользу альтернативных гипотез, то, пожалуй, единственным преимуществом функционала (4) остается его простота.

Для получения более устойчивой к аномальным наблюдениям \tilde{Y}_i оценки, например, можно использовать метод наименьших модулей [12]. При наличии базы данных по значениям работы выхода анализируемого образца [13] для построения $\hat{\vartheta}$ имеет смысл опираться на эту априорную информацию. Поиск минимума нового функционала в таких случаях может осложниться ввиду потери гладкости по параметрам.

Цель данной работы — анализ особенностей применения метода случайного поиска для оценки параметров сигнала ПЭЭ.

2. Математическая модель сигнала. Рассмотрим задачи, необходимые для достижения поставленной цели, в рамках математического моделирования. Это позволит сравнить полученные результаты с данными, закладываемыми в модель.

Пусть наблюдения \tilde{I}_i получены при известных значениях параметров (q_1^*, q_2^*) с использованием закона (1). Погрешности ϵ будем считать случайной величиной, разыгрываемой по правилу

$$\epsilon_i = \delta \frac{I_i}{I_0} \epsilon'_i, \quad (6)$$

где безразмерная величина δ^2 отвечает за дисперсию наблюдений, ϵ'_i являются независимыми реализациями стандартной нормальной случайной величины [10]. В данном подходе интенсивность шума получается пропорциональной сигналу. Переход к координатам ФН дает линейаризованные наблюдения \tilde{Y}_i :

$$\tilde{Y}_i = \lg \left[\frac{\tilde{I}_i}{V^2} \frac{V_0^2}{I_0} \right] = \vartheta_1 + \vartheta_2 X_i + \frac{\epsilon'_i \delta}{\ln 10} + o(\epsilon'_i \delta). \quad (7)$$

Обратим внимание, что в модель заложены независимость и нормальность. Гомоскедастичность остатков (5) для линейаризованной регрессионной модели должна выполняться в пределе при $\delta \rightarrow 0$. Однако эти особенности шума (6) не гарантируют того, что отдельно взятая выборка пройдет тесты на принадлежность нормальному распределению, отсутствие автокорреляции, равенство дисперсий при различных значениях X_i . Иногда в статистическом эксперименте даже значимость регрессионной модели может оказаться под сомнением [5].

На рис. 1 приведены ВАХ с разным количеством наблюдений и различным уровнем шума. Значения напряжения рассматривались равноотстоящими. Для наглядности при каждой величине V_i смоделировано 50 повторяющихся измерений.

3. Метод случайного поиска с обучением. Для минимизации (4) сначала выделим некоторую конечную подобласть множества Q . В описываемом случае это был прямоугольник $\Pi = [a_1, b_1] \times [a_2, b_2] = [6, 8] \times [6, 8]$, не содержащий точку достижения экстремума и точку (q_1^*, q_2^*) (см. рис. 2, где начальная область поиска выделена сплошным серым цветом). Таким образом, при процедуре оптимизации имеем дело с функционалом, наследуемым от (4):

$$J(\xi; \mathbf{X}, \tilde{\mathbf{Y}}) = \sum_{i=1}^N \left(\tilde{Y}_i - \lg \xi_1 + \frac{\xi_2}{\ln 10} X_i \right)^2, \quad (4')$$

где $\xi = (\xi_1, \xi_2)^T$ — случайные испытания или пробы.

Далее идеология метода заключается в том, что экстремум следует искать интенсивнее в той части области Q , где уже достигнуты «хорошие» результаты методом

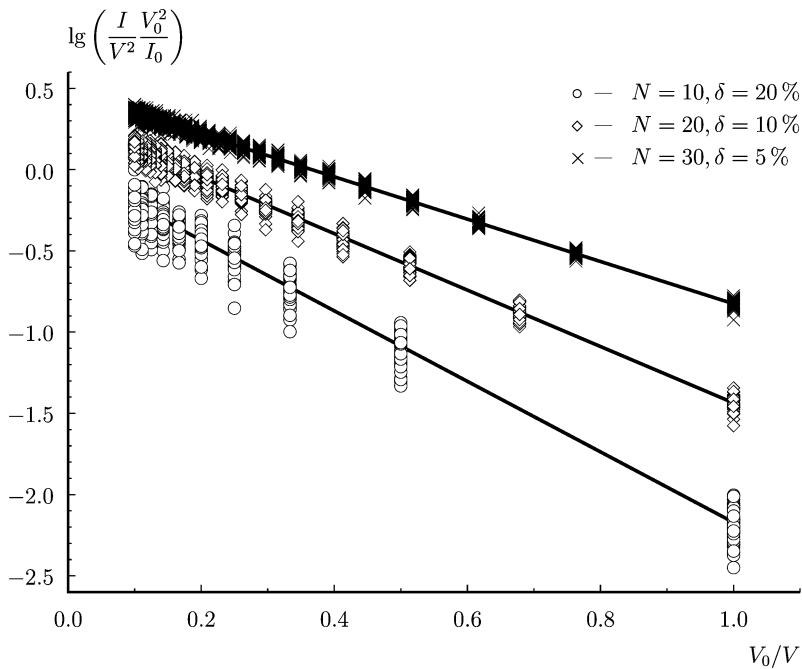


Рис. 1. Моделируемые ВАХ

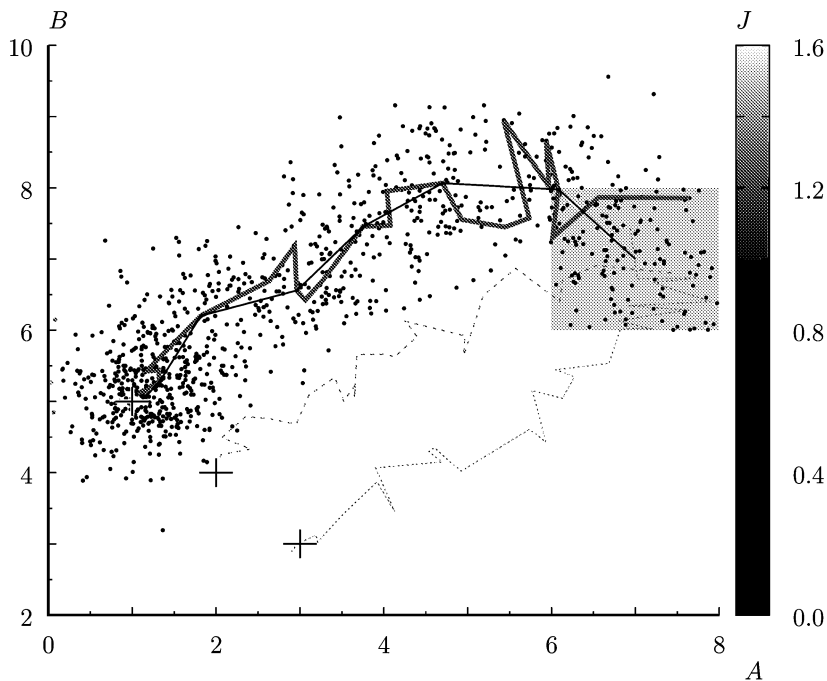


Рис. 2. Результаты поиска с обучением для $J(\theta; \mathbf{X}, \tilde{\mathbf{Y}})$

проб. С математической точки зрения это означает поиск решения с помощью плотности распределения, которая формируется на основе предыдущих опытов. В чем-то

такой подход напоминает метод существенной выборки при монте-карловском вычислении интеграла [14]. Для построения необходимой плотности используют оптимизируемую характеристику (4'). Поскольку

$$J(\xi; \mathbf{X}, \tilde{\mathbf{Y}}) \xrightarrow{\|\xi\| \rightarrow +\infty} +\infty, \quad J(\xi; \mathbf{X}, \tilde{\mathbf{Y}}) \xrightarrow{\delta \rightarrow 0} 0,$$

было решено отказаться от построения вспомогательного функционала с помощью традиционных рекомендаций: $L(\xi; \mathbf{X}, \tilde{\mathbf{Y}}) = -J(\xi; \mathbf{X}, \tilde{\mathbf{Y}})$ или $L(\xi; \mathbf{X}, \tilde{\mathbf{Y}}) = 1/J(\xi; \mathbf{X}, \tilde{\mathbf{Y}})$. В данном случае производился поиск максимума выражения

$$L(\xi) = L(\xi; \mathbf{X}, \tilde{\mathbf{Y}}) = \frac{1}{1 + J(\xi; \mathbf{X}, \tilde{\mathbf{Y}})}, \quad (4'')$$

область возможных значений которого представляет собой интервал $(0, 1]$.

Проведем равномерный поиск в области Π : $\xi = (\xi_1, \xi_2)$, где $\xi_j^{(k)} = a_j + (b_j - a_j)\gamma_j^{(k)}$, $j = \overline{1, p}$, $k = \overline{1, N_{\text{try}}}$, $\gamma_j^{(k)}$ являются реализациями равномерно распределенной случайной (псевдослучайной) величины в интервале $[0, 1)$. Решение формируется последовательно:

$$\begin{aligned} L^{(1)} &= L(\xi^{(1)}), & \mathbf{q}^{(1)} &= \xi^{(1)}, \\ L^{(k)} &= \begin{cases} L(\xi^{(k)}), & L(\xi^{(k)}) > L^{(k-1)}, \\ L^{(k-1)}, & L(\xi^{(k)}) \leq L^{(k-1)}, \end{cases} & \mathbf{q}^{(k)} &= \begin{cases} \xi^{(k)}, & L(\xi^{(k)}) > L^{(k-1)}, \\ \mathbf{q}^{(k-1)}, & L(\xi^{(k)}) \leq L^{(k-1)}, \end{cases} \\ & & k &= \overline{2, N_{\text{try}}}, & \hat{\mathbf{q}} &= \mathbf{q}^{(N_{\text{try}})}. \end{aligned} \quad (8)$$

На основе N_{try} попыток строится ковариационная матрица \mathbf{C} с компонентами

$$c_{jl} = \frac{\sum_{k=1}^{N_{\text{try}}} \xi_j^{(k)} \xi_l^{(k)} L(\xi^{(k)})}{\sum_{k=1}^{N_{\text{try}}} L(\xi^{(k)})} - \frac{\sum_{k=1}^{N_{\text{try}}} \xi_j^{(k)} L(\xi^{(k)})}{\sum_{k=1}^{N_{\text{try}}} L(\xi^{(k)})} \frac{\sum_{k=1}^{N_{\text{try}}} \xi_l^{(k)} L(\xi^{(k)})}{\sum_{k=1}^{N_{\text{try}}} L(\xi^{(k)})}, \quad j, l = \overline{1, p}. \quad (9)$$

Для $\hat{\mathbf{q}}$ и \mathbf{C} существует двумерное гауссовское распределение нового вектора ξ такое, что его плотность

$$p_{\xi}(\mathbf{q}) = \frac{1}{\sqrt{(2\pi)^p \det \mathbf{C}}} \exp \left[-\frac{1}{2} (\mathbf{q} - \hat{\mathbf{q}})^T \mathbf{C}^{-1} (\mathbf{q} - \hat{\mathbf{q}}) \right].$$

Таким образом, далее поиск проводится во всей области Q с помощью $p_{\xi}(\mathbf{q})$, а точки ξ , не попавшие в нее, отбрасываются — происходит так называемое усечение распределения [14]. Это обстоятельство позволяет рассматривать многоэкстремальные задачи, поскольку всегда существует отличная от нуля вероятность оказаться в окрестности точки глобального максимума или минимума. Ввиду того, что матрица \mathbf{C} симметрична, существует нижняя треугольная матрица \mathbf{T} такая, что $\mathbf{C} = \mathbf{T}\mathbf{T}^T$. Это позволяет моделировать ξ по формуле $\xi = \mathbf{T}\zeta + \hat{\mathbf{q}}$, где компоненты ζ являются независимыми стандартными нормальными случайными величинами [15]. Элементы матрицы \mathbf{T} при этом находятся последовательно из системы уравнений

$$\sum_{k=1}^l t_{jk} t_{lk} = c_{jl}, \quad 1 \leq l \leq j \leq p.$$

Теперь можно последовательно получать обновляемую оценку $\hat{\mathbf{q}}$, согласно (8), и построить уточняемую ковариационную матрицу \mathbf{C} в соответствии с (9). Выбор количества таких итераций $N_{\text{iter}} \geq 2$, равно как и объем статистических испытаний N_{try} на каждой стадии, остается за исследователем. В рассматриваемом случае $N_{\text{iter}} = 10$, $N_{\text{try}} = 100$, т. е. на поиск одной окончательной оценки $\hat{\mathbf{q}}$ тратилось $N_{\text{search}} = 1000$ испытаний, что очень немного для методов статистического моделирования. Весь процесс отображен на рис. 2 для наиболее зашумленного сигнала, где точками указаны пробные испытания, интенсивность цвета которых соответствует значениям функционала (4'). Серой толстой линией соединены все промежуточные экстремумы, черной — центры поэтапных распределений ($\hat{\mathbf{q}}$), начиная с центра начальной области поиска — прямоугольника Π , серыми пунктирными — промежуточные минимумы (4') (максимумы (4'')) для двух оставшихся сигналов, крестиками отмечены заложенные в модель сигнала точные значения параметров. Показатели минимизируемого функционала показаны на рис. 3.

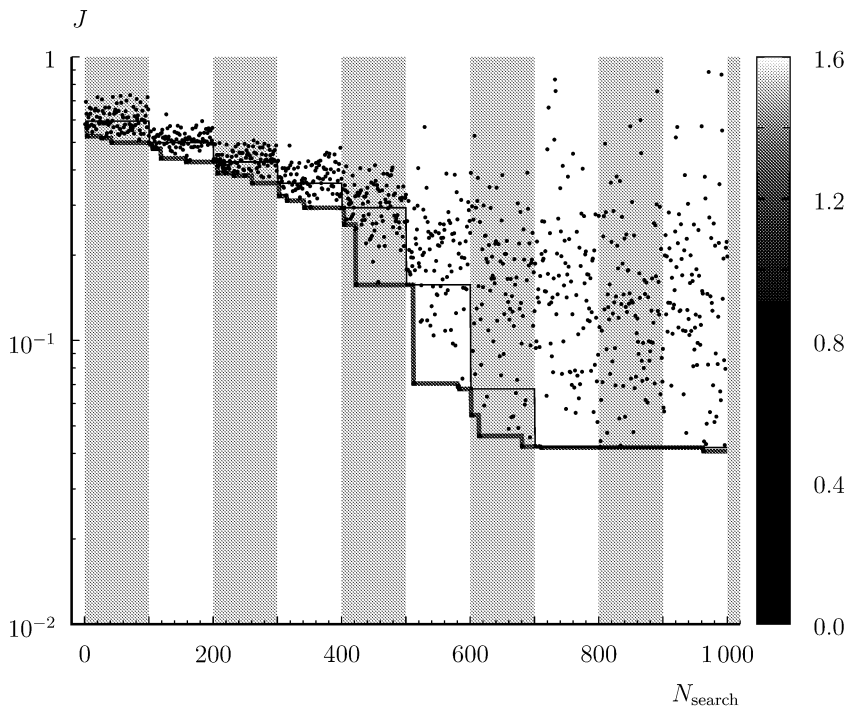


Рис. 3. Значения функционала $J(\vartheta; \mathbf{X}, \tilde{\mathbf{Y}})$

4. Результаты и статистические выводы. В рамках математического моделирования удобно сравнить результаты, полученные двумя разными способами: МНК-оценки \hat{q} ($\hat{\vartheta}$) и оценки случайного поиска \hat{q} ($\hat{\vartheta}$). Был проведен статистический эксперимент с $N_{\text{stat}} = 1000$ сигналами при каждом рассматриваемом уровне шума. Для демонстрации распределения оценок исходных значений параметров удобно использовать величины ϑ , поскольку в этом случае область рассеяния имеет эллиптическую форму. Данный тезис подтверждается тем, что совместная доверительная область для ϑ при выполнении предположений относительно остатков регрессионной

модели является внутренностью эллипса:

$$(\vartheta - \hat{\vartheta})^T \mathbf{M}(\vartheta - \hat{\vartheta}) \leq ps^2 F_{1-\alpha}(k_1, k_2), \quad s^2 = \frac{1}{k_2} \sum_{i=1}^N \hat{\epsilon}_i^2, \quad k_1 = p - 1, \quad k_2 = N - p;$$

$F_{1-\alpha}(k_1, k_2)$ — $(1 - \alpha)$ -квантиль F -распределения Фишера—Снедекора [10]. Для построения единого семейства эллипсов его центр был заменен на ϑ^* , а s^2 — на главный член дисперсии погрешности после линеаризации, согласно (7):

$$(\vartheta - \vartheta^*)^T \mathbf{M}(\vartheta - \vartheta^*) \leq p \frac{N}{k_2} \left(\frac{\delta}{\ln 10} \right)^2 F_{1-\alpha}(k_1, k_2).$$

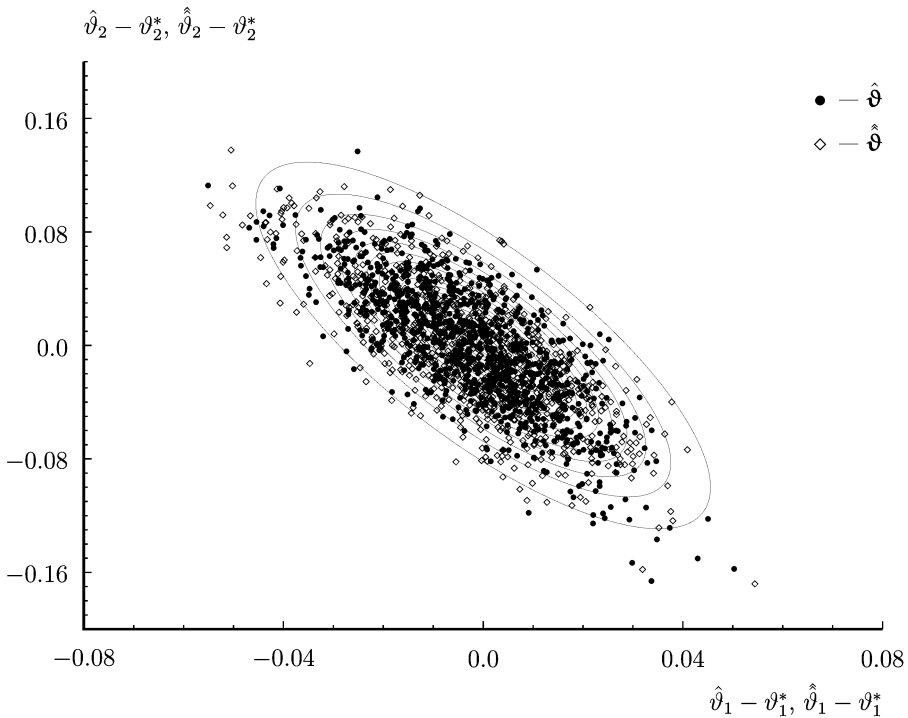


Рис. 4. Разбиение области Θ и рассеяние оценок

На рис. 4 представлены эллипсы со значениями α от 5 до 95 % с шагом 5 %. Таким образом, область Θ была разбита на $r = 20$ частей, которые были перенумерованы. Здесь же приведены оценки $\hat{\vartheta}$ и $\hat{\hat{\vartheta}}$ для $\delta = 10$ %. Предварительный визуальный анализ рассеяния параметров (имеется 1000 оценок каждого вида) позволяет выдвинуть гипотезу H_0 , согласно которой случайные величины $\hat{\vartheta}$ и $\hat{\hat{\vartheta}}$ подчиняются одному закону распределения. Можно записать, что

$$\hat{\vartheta} = \vartheta^* + \Delta_\varepsilon \vartheta, \quad \hat{\hat{\vartheta}} = \vartheta^* + \Delta_\varepsilon \vartheta + \Delta_{N_{\text{search}}} \vartheta. \quad (10)$$

Тогда H_0 будет выглядеть как $\|\Delta_{N_{\text{search}}} \vartheta\| \ll \|\Delta_\varepsilon \vartheta\|$. Это не вполне четкая запись с точки зрения математической статистики. По сути, перед нами гипотеза об однородности многомерных выборок. В описываемом случае компоненты векторов $\hat{\vartheta}$ и $\hat{\hat{\vartheta}}$

не являются независимыми. Построим вспомогательные одномерные случайные величины $\hat{\theta}$ и $\hat{\hat{\theta}}$. Пусть величина $\hat{\theta}$ принимает значение, равное номеру подобласти Θ , в которую попала оценка $\hat{\vartheta}$. Для $\hat{\hat{\theta}}$ аналогично. Частоты $(\hat{\nu}_k, \hat{\hat{\nu}}_k)$ попаданий преобразованных случайных величин в конкретную область и кумулятивные суммы частот $(\hat{F}_k, \hat{\hat{F}}_k)$ для случая $\delta = 10\%$ иллюстрирует рис. 5. Значения \hat{F}_k и $\hat{\hat{F}}_k$ описывают выборочные функции распределения.

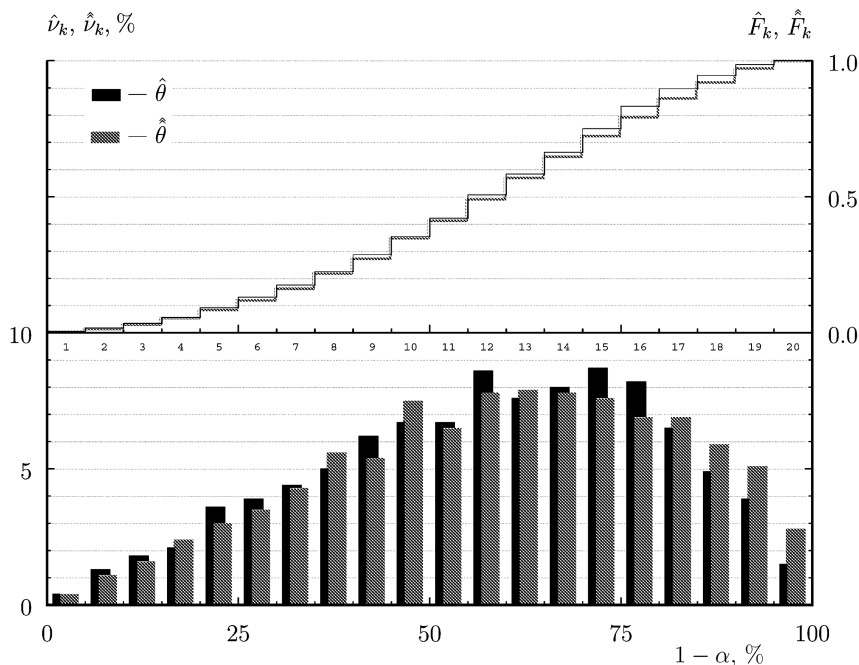


Рис. 5. Индивидуальные и накопленные частоты

Выдвинем вспомогательную гипотезу H'_0 , согласно которой выборки $\hat{\theta}$ и $\hat{\hat{\theta}}$ извлекаются из одной генеральной совокупности:

$$P\{\hat{\theta} \leq a\} = P\{\hat{\hat{\theta}} \leq a\} \quad \forall a.$$

Для ее проверки осуществим свободный от распределения тест Колмогорова — Смирнова [16]. Итак, имеются наблюдения $\hat{\theta}_1, \dots, \hat{\theta}_m$ и $\hat{\hat{\theta}}_1, \dots, \hat{\hat{\theta}}_m$; $m = 1000$. Обе выборки построены на основе независимых случайных векторов, извлеченных из непрерывных совокупностей. Далее выборки объединяются и упорядочиваются по возрастанию. Таким образом, строится вариационный ряд $Z_{(1)} \leq Z_{(2)} \leq \dots \leq Z_{(2m)}$. Это позволяет упростить процедуру построения гистограмм и выборочных функций распределения (рис. 5). Исследуется статистика

$$K_m = \sqrt{\frac{m}{2}} \max_{k=1, \dots, r} \{|\hat{F}_k - \hat{\hat{F}}_k|\}.$$

При $m \rightarrow \infty$ величина K_m подчиняется распределению с функцией Колмогорова. Если $K_m < K(\alpha)$, то основания отвергать H'_0 отсутствуют. Для 5%-го уровня значимости $K(\alpha) = 1.358$ (соответствующую величину можно найти в таблицах [16]).

Значения статистики для показателей шума 5, 10 и 20 % оказались равными 1.006, 0.8721 и 0.6037 соответственно. В рамках представления (10) поведение наблюдаемых (оцениваемых по выборочным данным) значений K_m означает, что при рассматриваемых величинах δ вкладом $\Delta_{N_{\text{search}}}$ можно пренебречь. При понижении уровня шума, оставляя N_{search} постоянным, рано или поздно будем вынуждены отклонить гипотезу H'_0 .

Предположение H'_0 можно проверить, используя метрику критерия ω^2 (Крамера — фон Мизеса — Смирнова). Это было предложено в работе Леманна [17] и проанализировано Розенблаттом [18]. Практически пригодная статистика приведена Андерсоном [19]:

$$T_m = \frac{1}{2m^2} \left[\sum_{i=1}^m (\hat{r}_i - i)^2 + \sum_{i=1}^m (\hat{r}_i - i)^2 \right] - \frac{4m^2 - 1}{12m},$$

где \hat{r}_i и \hat{r}_i — ранги представителей выборок $\hat{\theta}$ и $\hat{\theta}$ в общем вариационном ряду $\{Z_{(i)}\}$. При наличии совпадающих наблюдений следует использовать средние ранги.

При $m \rightarrow \infty$ величина T_m обладает функцией распределения $a_1(x)$ (см., например, [16]). Критическая область для 5%-го уровня значимости определяется величиной $T(0.05) = 0.4614$, оказавшейся больше, чем наблюдаемые значения 0.2428, 0.1798 и 0.1145 статистики T_m . Они приведены по-прежнему в порядке возрастания уровня шума. Оснований отвергать вспомогательные гипотезы однородности нет. Здесь наблюдается та же тенденция в динамике значения T_m , что и при использовании теста Колмогорова — Смирнова.

Также гипотеза H'_0 была проверена при помощи статистики Пирсона

$$\chi_m^2 = m \sum_{k=1}^r \frac{(\hat{\nu}_k - \hat{\nu}_k)^2}{\hat{\nu}_k + \hat{\nu}_k}$$

с количеством степеней свободы, равным $(r - 1)$, которая при $m \rightarrow \infty$ имеет соответствующее распределение [20]. Критическая точка при той же строгости к уровню доверия $\chi^2(0.05; 19) = 30.14$. Это значение больше, чем приведенные в той же последовательности, что и ранее, наблюдаемые величины 14.79, 11.54 и 20.84. Оснований отвергать гипотезы H'_0 опять не нашлось. Монотонная зависимость χ_m^2 от δ уже отсутствует. Это можно объяснить тем, что объектом анализа здесь являются отдельные, а не накопленные частоты. Можно отметить, что данный критерий не связан ограничениями на одномерность. Однако проведенное разбиение области Θ на ячейки с помощью эллипсов, по нашему мнению, наиболее логичное в таком случае.

5. Заключение. Была описана тестовая задача идентификации параметров сигнала ПЭЭ с помощью регрессионной модели, основанной на законе Фаулера — Нордгейма. Исследование проведено в рамках компьютерного статистического эксперимента. Методом, альтернативным МНК-оценке параметров, послужил случайный поиск с обучением. Установлено, что при рассматриваемых соотношениях уровня шума к мощности сигнала погрешностью случайного поиска можно до некоторой степени пренебречь при количестве статистических испытаний, равном $N_{\text{search}} = 1000$. Речь идет о рассеянии оцениваемых параметров относительно их точных значений, которое всегда имеет место. Такое заключение позволяет расширить класс используемых для идентификации отклика функционалов, не меняя при этом метода. Данный статистический подход может применяться и в натурном эксперименте с автоматизированными системами сбора информации [3]. Ничтожное для методов Монте-Карло

значение N_{search} позволяет получать оценки достаточно быстро. Стоит указать и на традиционные положительные особенности случайного поиска: сравнительно простое распараллеливание программного кода и работоспособность при наличии локальных экстремумов. Метод также может применяться в режиме реального времени, когда вычисления требуется прервать и выдать в качестве результата лучшую из имеющихся проб. При этом совместные доверительные области распределения параметров сигнала являются одними из важнейших показателей качества проведенного регрессионного анализа [5]. В рассматриваемой тестовой задаче эти множества определяются семейством эллипсов. Однако в общем случае их форма зависит от используемого функционала [21]. И поскольку есть основания считать, что метод случайного поиска практически не оказывает влияния на рассеяние параметров, благодаря ему можно построить доверительные области любой сложности с помощью, например, независимой от распределения процедуры [22].

Литература

1. Бугаев А. С., Виноградова Е. М., Егоров Н. В., Шешин Е. П. Автоэлектронные катоды и пушки. Долгопрудный: Издат. Дом «Интеллект», 2017. 288 с.
2. Vinogradova E. M., Starikova A. V., Varayun' M. I. Multipole electrostatic system mathematical modeling // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2017. Т. 13. Вып. 4. С. 365–371. <https://doi.org/10.21638/11701/spbu10.2017.403>
3. Егоров Н. В., Шешин Е. П. Автоэлектронная эмиссия. Принципы и приборы. Долгопрудный: Издат. Дом «Интеллект», 2011. 704 с.
4. Фоменко В. С. Эмиссионные свойства материалов: справочник. Изд. 4-е, перераб. и доп. Киев: Наукова Думка, 1981. 340 с.
5. Egorov N. V., Antonov A. Yu., Varayun' M. I. Analysis of confidence intervals for regression model parameters, based on the Fowler – Nordheim law // Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques. 2020. Vol. 14. N 4. P. 730–737.
6. Fowler R. H., Nordheim L. W. Electron emission in intense electric fields // Proceedings of the Royal Society A. Mathematical, Physical and Engineering Sciences. 1928. Vol. 119. N 781. P. 173–181.
7. Антонов А. Ю., Варяюнь М. И., Егоров Н. В. Линеаризованная трехпараметрическая регрессионная модель для сигнала полевой электронной эмиссии // Нано- и микросистемная техника. 2019. Т. 21. № 2. С. 103–110.
8. Egorov N. V., Varayun' M. I., Bure V. M., Antonov A. Yu. Regression models for the field electron emission signal // Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques. 2020. Vol. 14. N 6. P. 1394–1402.
9. Egorov N. V., Antonov A. Yu., Gribkova I. M. Statistical test of a single semiempirical work function model // Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques. 2014. Vol. 8. P. 138–143.
10. Egorov N. V., Antonov A. Yu., Varayun' M. I. Analysis of the emission characteristics of field cathodes using regression models // Journal of Surface Investigation: X-Ray, Synchrotron and Neutron Techniques. 2018. Vol. 12. P. 1005–1012.
11. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. В 2-х кн. Кн. 1. Изд. 2-е, перераб. и доп. / пер. с англ. М.: Финансы и статистика, 1986. 366 с.
12. Буре В. М., Париллина Е. М., Седаков А. А. Методы прикладной статистики в R и Excel. СПб.: Лань, 2016. 152 с.
13. Antonov A. Yu., Varayun' M. I., Gribkova I. M., Pigul' E. Yu. Mathematical modelling of the work function distribution on a monocrystalline cathode surface // 2015 International Conference “Stability and Control Processes” in memory of V. I. Zubov (SCP). St. Petersburg: Publishing House Fedorova G. V., 2015. P. 144–147.
14. Владимирова Л. В., Овсянников Д. А., Рубцова И. Д. Методы Монте-Карло в прикладных задачах. СПб.: ВВМ, 2015. 167 с.
15. Соболев И. М. Численные методы Монте-Карло. М.: Наука, 1973. 312 с.
16. Большев Л. Н., Смирнов Н. В. Таблицы математической статистики. М.: Наука. Гл. ред. физ.-мат. лит., 1983. 416 с.
17. Lehmann E. L. Consistency and unbiasedness of certain nonparametric tests // The Annals of Mathematical Statistics. 1951. Vol. 22. N 2. P. 165–179.

18. Rosenblatt M. Limit theorems associated with variants of the von Mises statistic // The Annals of Mathematical Statistics. 1952. Vol. 23. N 4. P. 617–623.

19. Anderson T. W. On the distribution of the two-sample Cramér — von Mises criterion // The Annals of Mathematical Statistics. 1962. Vol. 33. N 3. P. 1148–1159.

20. Тюрин Ю. Н., Макаров А. А. Анализ данных на компьютере / под ред. В.Э. Фигурнова. 3-е изд., перераб. и доп. М.: ИНФРА-М, 2003. 544 с.

21. Прохоров И. Д., Суртаева М. Н., Антонов А. Ю. Использование равномерных последовательностей для параметрической идентификации сигнала полевой электронной эмиссии // Процессы управления и устойчивость. 2018. Т. 5. № 1. С. 191–195.

22. Hu Zh., Yang R.-C. A new distribution-free approach to constructing the confidence region for multiple parameters // PLoS One. 2013. Vol. 8. N 12. P. 1–13.

Статья поступила в редакцию 18 марта 2021 г.

Статья принята к печати 4 июня 2021 г.

Контактная информация:

Вараюнь Марина Ивановна — канд. физ.-мат. наук, доц.; m.varayun@spbu.ru

Виноградова Екатерина Михайловна — д-р физ.-мат. наук, проф.; e.m.vinogradova@spbu.ru

Антонов Андрей Юрьевич — канд. физ.-мат. наук, доц.; a.antonov@spbu.ru

Random search method for estimating the parameters of the emission system signal*

M. I. Varayun', E. M. Vinogradova, A. Yu. Antonov

St. Petersburg State University, 7–9, Universitetskaya nab., St. Petersburg, 199034, Russian Federation

For citation: Varayun' M. I., Vinogradova E. M., Antonov A. Yu. Random search method for estimating the parameters of the emission system signal. *Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes*, 2021, vol. 17, iss. 3, pp. 228–239. <https://doi.org/10.21638/11701/spbu10.2021.301> (In Russian)

Within the framework of a computer statistical experiment, the test problem of identifying the parameters of a field electron emission signal using a regression model based on the Fowler — Nordheim law is considered. Two approaches to determining the parameter estimates are used in the work: ordinary least squares and random search with training. It is shown that the random search error can be neglected if, for the considered ratios of the noise level to the signal power, the number of statistical tests is approximately 10^3 . The result allows us to expand the class of functionals used to identify the response without changing the method. The article notes the advantages of the random search method for the problem under consideration and the prospects of its applicability to tasks in a more general setting.

Keywords: field electron emission, current-voltage characteristic, regression model, least-squares method, random search method.

References

1. Bugaev A. S., Vinogradova E. M., Egorov N. V., Sheshin E. P. *Avtojelektronnye katody i pushki* [Autoelectronic cathodes and guns]. Dolgoprudny, "Intellekt" Publ., 2017, 288 p. (In Russian)

2. Vinogradova E. M., Starikova A. V., Varayun' M. I. Multipole electrostatic system mathematical modeling. *Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes*, 2017, vol. 13, iss. 4, pp. 365–371. <https://doi.org/10.21638/11701/spbu10.2017.403>

3. Egorov N. V., Sheshin E. P. *Avtojelektronnaja jemissija. Principy i pribory* [Field emission. Principles and Instruments]. Dolgoprudny, "Intellekt" Publ., 2011, 704 p. (In Russian)

* This work was supported by the Russian Foundation for Basic Research (project N 20-07-01086).

4. Fomenko V. S. *Jemissionnyje svojstva materialov*. Spravochnik [*Emission properties of materials*. Reference book]. 4th ed. Kiev, Naukova Dumka Publ., 1981, 340 p. (In Russian)
5. Egorov N. V., Antonov A. Yu., Varayun' M. I. Analysis of confidence intervals for regression model parameters, based on the Fowler — Nordheim law. *Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques*, 2020, vol. 14, no. 4, pp. 730–737.
6. Fowler R. H., Nordheim L. W. Electron emission in intense electric fields. *Proceedings of the Royal Society A. Mathematical, Physical and Engineering Sciences*, 1928, vol. 119, no. 781, pp. 173–181.
7. Antonov A. Yu., Varayun' M. I., Egorov N. V. Linearizovannaja trjohparametricheskaja regressionnaja model' dlja signala polevoj jelektronnoj jemissii [Linearized three-parameter regression model for the field electron emission signal]. *Nano- and Microsystems Technology*, 2019, vol. 21, no. 2, pp. 103–110. (In Russian)
8. Egorov N. V., Varayun' M. I., Bure V. M., Antonov A. Yu. Regression models for the field electron emission signal. *Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques*, 2020, vol. 14, no. 6, pp. 1394–1402.
9. Egorov N. V., Antonov A. Yu., Gribkova I. M. Statistical test of a single semiempirical work function model. *Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques*, 2014, vol. 8, pp. 138–143.
10. Egorov N. V., Antonov A. Yu., Varayun' M. I. Analysis of the emission characteristics of field cathodes using regression models. *Journal of Surface Investigation: X-Ray, Synchrotron and Neutron Techniques*, 2018, vol. 12, pp. 1005–1012.
11. Draper N., Smith H. *Applied regression analysis*. 2nd ed. New York, Wiley Publ., 1981, 736 p. (Rus. ed.: Draper N., Smith H. *Prikladnoj regressionnyj analiz*. In 2 books. Book 1. 2nd ed. Moscow, Finansy i statistika Publ., 1986, 366 p.)
12. Bure V. M., Parilina E. M., Sedakov A. A. *Metody prikladnoj statistiki v R i Excel* [*Methods of applied statistics in R and Excel*]. St. Petersburg, Lan' Publ., 2016, 152 p. (In Russian)
13. Antonov A. Yu., Varayun' M. I., Gribkova I. M., Pigul' E. Yu. Mathematical modelling of the work function distribution on a monocrystalline cathode surface. *2015 International Conference "Stability and Control Processes" in memory of V. I. Zubov (SCP)*. St. Petersburg, Publ. House Fedorova G. V., 2015, pp. 144–147.
14. Vladimirova L. V., Ovsjannikov D. A., Rubcova I. D. *Metody Monte-Karlo v prikladnyh zadachah* [*The Monte Carlo method in applied tasks*]. St. Petersburg, VVM Publ., 2015, 167 p. (In Russian)
15. Sobol' I. M. *Chislennye metody Monte-Karlo* [*Numerical Monte Carlo methods*]. Moscow, Nauka Publ., 1973, 312 p. (In Russian)
16. Bolshev L. N., Smirnov N. V. *Tablicy matematicheskoj statistiki* [*Tables of mathematical statistics*]. Moscow, Nauka Publ., 1983, 416 p. (In Russian)
17. Lehmann E. L. Consistency and unbiasedness of certain nonparametric tests. *The Annals of Mathematical Statistics*, 1951, vol. 22, no. 2, pp. 165–179.
18. Rosenblatt M. Limit theorems associated with variants of the von Mises statistic. *The Annals of Mathematical Statistics*, 1952, vol. 23, no. 4, pp. 617–623.
19. Anderson T. W. On the distribution of the two-sample Cramér — von Mises criterion. *The Annals of Mathematical Statistics*, 1962, vol. 33, no. 3, pp. 1148–1159.
20. Tyurin Yu. N., Makarov A. A. *Analiz dannyh na komp'jutere* [*Data analysis on the computer*]. Ed. by V. E. Figurnov. Moscow, INFRA-M Publ., 2003, 544 p. (In Russian)
21. Prohorov I. D., Surtaeva M. N., Antonov A. Yu. Ispol'zovanie ravnomernyh posledovatel'nostej dlja parametricheskoj identifikacii signala polevoj jelektronnoj jemissii [Use of uniform sequences for the parametric identification of the field electron emission signal]. *Control Processes and Stability*, 2018, vol. 5, no. 1, pp. 191–195. (In Russian)
22. Hu Zh., Yang R. C. A new distribution-free approach to constructing the confidence region for multiple parameters. *PLoS One*, 2013, vol. 8, no. 12, pp. 1–13.

Received: March 18, 2021.

Accepted: June 04, 2021.

A u t h o r s ' i n f o r m a t i o n :

Marina I. Varayun' — PhD in Physics and Mathematics, Associate Professor; m.varayuan@spbu.ru

Ekaterina M. Vinogradova — Dr. Sci. in Physics and Mathematics, Professor;
e.m.vinogradova@spbu.ru

Andrei Yu. Antonov — PhD in Physics and Mathematics, Associate Professor; a.antonov@spbu.ru