Ein Axiomensystem der Methode der kleinsten Quadrate.

Von

F. Bernstein und W. S. Baer in Göttingen.

§ 1.
Einleitung.

Die erste wissenschaftliche Begründung des von Legendre*) unter dem Namen „Methode der kleinsten Quadrate" in die Ausgleichungsrechnung eingeführten Rechenverfahrens hat C. F. Gauß**) gegeben; bei dieser seiner ersten Ableitung benutzt er das aus seiner Hypothese des arithmetischen Mittels abgeleitete *Fehlergesetz*

$$\varphi(x) = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2},$$

dessen Integral über ein Intervall a bis b die Wahrscheinlichkeit für das Zustandekommen eines Fehlers x zwischen den Grenzen a und b angibt. Weiter hat Laplace***) aus der Annahme, daß sich jeder zu begehende Fehler aus einer großen Anzahl sehr kleiner sog. Elementarfehler zusammensetze, das Fehlergesetz herzuleiten gesucht und hat, gestützt auf dieses, eine Begründung der Methode der kleinsten Quadrate gegeben; seine Ableitung des Fehlergesetzes ist dann im wesentlichen von Bienaymé, Liapounoff, Markoff, Tschebyscheff †) in eine endgültige Form gebracht worden. Ferner hat C. F. Gauß ††) die Methode der kleinsten Quadrate

*) Nouvelles méthodes pour la détermination des orbites des comètes, Paris 1805—1806.

**) Theoria motus corporum coelestium etc, Hamburg 1809, Art. 186.

***) Théorie analytique des probabilités, Paris 1812, Art. 20—21.

†) Vgl. z. B. das Literaturverzeichnis am Ende der von A. Markoff herausgegebenen Festschrift „Bicentenaire de la loi des grands nombres 1713—1913, Démonstration du second théorème limite du calcul des probabilités par la méthode des moments", St. Petersburg 1913.

††) Theoria combinationis observationum erroribus minimis obnoxiae, Comm. Goett. I, 1821—1826.

auch noch auf eine zweite Weise zu begründen versucht, ohne diesmal auf den obigen Ausdruck des Fehlergesetzes zurückzugehen; diese klassische Beweismethode, die in viele Lehrbücher*) der Wahrscheinlichkeitsrechnung übergegangen ist, benutzt, indem sie den Begriff des *Fehlerrisikos* in den Mittelpunkt stellt, außer den Grundbegriffen und Grundsätzen der Wahrscheinlichkeitsrechnung nur die Existenz und einige allgemeine Eigenschaften einer Fehlerfunktion. Diese wahrscheinlichkeitstheoretischen Begriffe gleichfalls zu eliminieren, ist bisher unseres Wissens noch nicht einwandfrei gelungen**). Es wird nun im folgenden der Versuch gemacht, die Methode der kleinsten Quadrate unter Benutzung des von Gauß in seiner zweiten Begründung mit Erfolg verwandten Gewichtsbegriffes und *unter Vermeidung der Begriffe der Wahrscheinlichkeitsrechnung* axiomatisch zu begründen, wobei die in § 2 aufgeführten Axiome möglichst nach dem Gesichtspunkte ausgewählt sind, daß sie dem inneren Charakter des Ausgleichungsproblems entsprechen. Nachdem in § 3 einige mathematische Hilfssätze über reelle Funktionen zusammengestellt worden sind, werden in § 4 aus dem Axiomensystem die Folgerungen gezogen, die zu den in der Methode der kleinsten Quadrate üblichen Relationen unmittelbar hinführen.

Die Vorteile dieser Begründung liegen darin, daß zunächst die prinzipiellen Schwierigkeiten vermieden werden, die bei der Einführung der Gleichberechtigung von Fällen in der Wahrscheinlichkeitsrechnung jedesmal auftreten; ferner ist die Begründung unabhängig von jeder bisher benutzten speziellen Eigenschaft der Fehlerfunktion, wodurch die Tragweite der Methode wesentlich erweitert wird; dazu kommt, daß die benutzten Regeln für den Gebrauch des Gewichtsbegriffes einfacher als die Regeln für die Verwendung des Wahrscheinlichkeitsbegriffes ausfallen. Was die philosophische Kategorie des Gewichtsbegriffes angeht, so ist er offenbar dem Wahrscheinlichkeitsbegriff verwandt; es hängt aber ganz von dem gewählten Ausgangspunkte ab, welche Begriffsbildung man als die primitivere ansehen wird.

---

*) Es werde insbesondere das Lehrbuch von A. A. Markoff, Wahrscheinlichkeitsrechnung, deutsch von H. Liebmann, Leipzig-Berlin 1912, zitiert.

**) Die kürzlich erschienene Untersuchung von R. Suppantschitsch „Zur Axiomatik der Methode der kleinsten Quadrate" [Sitzungsberichte der k. Akad. der Wissensch. in Wien, Math.-naturw. Klasse, 122, Abt. 2a (1913), S. 3—34] liegt in einer anderen Richtung.

§ 2.

Das Axiomensystem.

1. *Zuordnungsaxiom:* Jede Zuordnung irgendeines Wertes α zu irgendeiner festen Größe a derart, daß für jedes reelle λ auch $\lambda\alpha$ zu λa zugeordnet wird, heiße eine Annäherung und werde bezeichnet durch

$$a \sim \alpha, \quad \lambda a \sim \lambda\alpha.$$

2. *Gewichtsaxiom:* Jeder Annäherung werde als Gewicht eine *positive* Zahl, die alle möglichen Werte annehmen kann, zugeordnet und dies ausgedrückt durch

$$a \sim \alpha \; [p].$$

3. *Axiom der Unabhängigkeit vom Maßstabe:* Ist auf Grund des Zuordnungsaxioms

$$a \sim \alpha \; [p], \quad \lambda a \sim \lambda\alpha \; [p'],$$

so sei der Quotient $p':p$ lediglich von λ abhängig

$$p':p = f(\lambda),$$

wobei $f(\lambda)$ eine stetige Funktion von λ, aber nicht identisch 1 sein soll.

4. *Axiom der Unabhängigkeit von Annäherungen:* Zwei Annäherungen

$$a \sim \alpha \; [p], \quad b \sim \beta \; [q]$$

mögen voneinander unabhängig heißen, wenn folgende Bedingung erfüllt ist: Ist auf Grund des Zuordnungsaxioms für irgendwelche reellen Werte λ und μ

$$\lambda a \sim \lambda\alpha \; [p'], \quad \mu b \sim \mu\beta \; [q']$$

so bestimme sich das Gewicht r der „komponierten" Annäherung

$$\lambda a + \mu b \sim \lambda\alpha + \mu\beta \; [r]$$

eindeutig durch die Funktionalgleichung

$$\psi(r) = \psi(p') + \psi(q'),$$

wobei $\psi(x)$ eine reelle, eindeutige und stetige Funktion des Gewichtsargumentes allein sei.

5. *Axiom des arithmetischen Mittels:* Das arithmetische Mittel zweier unabhängigen Annäherungswerte für dieselbe feste Größe mit dem Gewicht 1 besitze als Annäherungswert für dieselbe feste Größe das Gewicht 2.

6. *Axiom der Größenordnung der Funktion $\psi(x)$:* Sind

$$a \sim \alpha_1 \; [1], \; a \sim \alpha_2 \; [1], \; \cdots, \; a \sim \alpha_n \; [1]$$

irgend n zu je zweien unabhängige Annäherungen für dieselbe feste Größe a je mit dem Gewichte 1, so habe das Gewicht r der „komponierten" Annäherung

$$na \sim \alpha_1 + \alpha_2 + \cdots + \alpha_n \; [r]$$

die Größenordnung $\frac{1}{n}$ in dem Sinne, daß bei wachsender Anzahl n der Quotient von r und $\frac{1}{n}$ von einer Stelle an, d. h. für alle $n > n_0$, oberhalb einer positiven und unterhalb einer endlichen Schranke verbleibe.

7. *Ausgleichungsaxiom:* Sind $a_1, a_2, \cdots, a_m$ irgend m feste Größen und liegen für n lineare Ausdrücke in ihnen irgend n unabhängige Annäherungen vor

$$\begin{array}{l} A_{11}a_1 + A_{21}a_2 + \cdots + A_{m1}a_m \sim l_1 \; [p_1], \\ A_{12}a_1 + A_{22}a_2 + \cdots + A_{m2}a_m \sim l_2 \; [p_2], \\ \cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots \\ A_{1n}a_1 + A_{2n}a_2 + \cdots + A_{mn}a_m \sim l_n \; [p_n], \end{array}$$

wo $n > m$ und die A_{ik} bekannte Koeffizienten seien, so sei für irgendeinen linearen Ausdruck in den $a_1, a_2, \cdots, a_m$ diejenige „Komposition" der obigen n Annäherungen mit n zu bestimmenden Multiplikatoren $\lambda_1, \lambda_2, \cdots, \lambda_n$, die auf der linken Seite diesen linearen Ausdruck in den $a_1, a_2, \cdots, a_m$ liefert, die *beste Annäherung (Ausgleichung)*, die das größte Gewicht besitzt.

§ 3.

Hilfssätze über reelle Funktionen.

Hilfssatz 1: *Der Funktionalgleichung*

$$cf(tx_1x_2) = f(tx_1)f(tx_2),$$

in der $c \neq 0$ und $t \neq 0$ gegebene reelle Konstante sind, genügen alle Funktionen

$$f(x) = c\left(\frac{x}{t}\right)^\gamma,$$

wo γ beliebig reell ist, und dies sind auch die sämtlichen für alle reellen oder auch nur für alle reell-positiven Argumente definierten reellen, eindeutigen und stetigen Funktionen, die diese Funktionalgleichung befriedigen und nicht identisch verschwinden.

Mit diesem bekannten Satze beweist man weiter, wie uns zuerst Herr S. Halberstadt mitgeteilt hat, den

Hilfssatz 2: *Die Gesamtheit der für alle $x > 0$ reellen, eindeutigen und* *stetigen Funktionen* $\psi(x)$, *bei denen es zu jedem Argumentenpaar* $x_1 > 0$, $x_2 > 0$ *stets genau ein Argument* $\xi = \xi(x_1, x_2) > 0$ *gibt, sodaß*

$$\psi(\xi) = \psi(x_1) + \psi(x_2) \tag{1}$$

gilt, und bei denen zugleich mit (1) *für alle reellen* $\varrho > 0$ *auch*

$$\psi(\varrho\xi) = \psi(\varrho x_1) + \psi(\varrho x_2) \tag{2}$$

erfüllt ist, wird geliefert durch

$$\psi(x) = kx^\gamma,$$

wo $k \neq 0$ *und* $\gamma \neq 0$ *beliebige reelle Konstanten bedeuten.*

Beweis: Zunächst verschwindet $\psi(x)$ für $x > 0$ nicht identisch, da sonst entgegen der Voraussetzung (1) nicht ξ durch x_1 und x_2 eindeutig bestimmt wäre; also gibt es mindestens ein Argument $x_0 > 0$, so daß

$$\psi(x_0) \neq 0 \tag{3}$$

ist. Weiter folgt durch mehrmalige Anwendung der Voraussetzungen (1) und (2), daß es allgemein für jede*) ganze Zahl $l \geqq 1$ zu irgend l positiven Argumenten $x_1, x_2, \cdots, x_l$ ein Argument $\xi = \xi(x_1, x_2, \cdots, x_l) > 0$ gibt, derart, daß für alle $\varrho > 0$

$$\psi(\varrho\xi) = \psi(\varrho x_1) + \psi(\varrho x_2) + \cdots + \psi(\varrho x_l)$$

gilt; $x_1 = x_2 = \cdots = x_l = x_0$ liefert für jede**) ganze positive Zahl l die Existenz eines Argumentes $\xi_l > 0$, so daß für alle $\varrho > 0$

$$\psi(\varrho\xi_l) = l\,\psi(\varrho x_0) \tag{4}$$

ist. Gehören auf diese Weise zu den ganzen positiven Zahlen m und n die Argumente $\xi_m > 0$ und $\xi_n > 0$, so ist also für alle $\varrho > 0$

$$\psi(\varrho\xi_m) = m\,\psi(\varrho x_0), \quad \psi(\varrho\xi_n) = n\,\psi(\varrho x_0)$$

insbesondere für $\varrho = 1$ und $\varrho = \frac{\xi_m}{\xi_n} > 0$

$$\psi(\xi_m) = m\,\psi(x_0), \quad \psi(\xi_m) = n\,\psi\left(\frac{\xi_m}{\xi_n} x_0\right);$$

hieraus folgt

$$\psi\left(\frac{\xi_m}{\xi_n} x_0\right) = \frac{m}{n}\,\psi(x_0). \tag{5}$$

Wegen (3) nimmt daher $\psi(x)$ für $x > 0$ alle rational-positiven Multipla von $\psi(x_0)$ und wegen der vorausgesetzten Stetigkeit alle reellen Werte an, die dasselbe Vorzeichen wie $\psi(x_0)$ haben.

Ferner nimmt $\psi(x)$ jeden solchen Wert genau einmal für $x > 0$ an; denn wäre

$$\psi(x') = \psi(x''), \qquad x' \neq x'',\ x' > 0,\ x'' > 0,$$

*) Für $l = 1$ ist insbesondere $\xi = x_1$.

**) Für $l = 1$ ist insbesondere $\xi_1 = x_0$ zu setzen.

so gäbe es nach (1) und (2) zu x' und x'' ein Argument $\xi' = \xi'(x', x'') > 0$, so daß für alle $\varrho > 0$

$$\psi(\varrho\xi') = \psi(\varrho x') + \psi(\varrho x'')$$

wäre, insbesondere für $\varrho = \frac{x'}{\xi'} > 0$

$$\psi(x') = \psi\left(\frac{x'^2}{\xi'}\right) + \psi\left(\frac{x'x''}{\xi'}\right);$$

es gäbe also entgegen der Voraussetzung (1) zu dem Argumentenpaar $\frac{x'^2}{\xi'} > 0$ und $\frac{x'x''}{\xi'} > 0$ zwei verschiedene Argumente $x' > 0$ und $x'' > 0$, für die

$$\psi(x') = \psi(x'') = \psi\left(\frac{x'^2}{\xi'}\right) + \psi\left(\frac{x'x''}{\xi'}\right)$$

wäre. Also ist $\psi(x)$ in der Tat einwertig, mithin nach dem Vorangehenden monoton und daher eine monoton-einwertige Funktion.

Weiter verschwindet $\psi(x)$ nirgends für $x > 0$; denn gesetzt, es wäre für $\bar{x} > 0$

$$\psi(\bar{x}) = 0,$$

so folgte aus (4) mit $l = 2$, $\varrho = \frac{\bar{x}}{x_0} > 0$

$$\psi\left(\frac{\bar{x}}{x_0}\xi_2\right) = 2\psi(\bar{x}) = 0 = \psi(\bar{x}),$$

und dies widerspräche wegen $\frac{\bar{x}}{x_0}\xi_2 \neq \bar{x}$ der eben bewiesenen Einwertigkeit von $\psi(x)$.

Endlich erfüllt $\psi(x)$ für alle $x_1 > 0$ und $x_2 > 0$ die Funktionalgleichung

$$\psi(x_0)\,\psi(x_0 x_1 x_2) = \psi(x_0 x_1)\,\psi(x_0 x_2). \tag{6}$$

Ist nämlich zunächst mindestens eines der beiden Argumente x_1 und x_2, etwa x_2, eines der vorher definierten $\frac{\xi_m}{\xi_n}$, für die (5) galt, so wird aus (4) mit $l = m$, $\varrho = \frac{x_0 x_1}{\xi_n} > 0$ sowie $l = n$, $\varrho = \frac{x_0 x_1}{\xi_n} > 0$

$$\psi\left(x_0 x_1 \frac{\xi_m}{\xi_n}\right) = m\,\psi\left(\frac{x_0^2 x_1}{\xi_n}\right), \quad \psi(x_0 x_1) = n\,\psi\left(\frac{x_0^2 x_1}{\xi_n}\right);$$

da, wie bewiesen, $\psi\left(\frac{x_0^2 x_1}{\xi_n}\right)$ wegen $\frac{x_0^2 x_1}{\xi_n} > 0$ nicht verschwindet, so folgt hieraus in Verbindung mit (5)

$$\frac{\psi\left(x_0 x_1 \frac{\xi_m}{\xi_n}\right)}{\psi(x_0 x_1)} = \frac{m}{n} = \frac{\psi\left(x_0 \frac{\xi_m}{\xi_n}\right)}{\psi(x_0)}$$

oder

$$\psi(x_0)\,\psi\left(x_0 x_1 \frac{\xi_m}{\xi_n}\right) = \psi(x_0 x_1)\,\psi\left(x_0 \frac{\xi_m}{\xi_n}\right), \tag{7}$$

womit (6) zunächst für $x_2 = \frac{\xi_m}{\xi_n}$ bewiesen ist. Allgemein läßt sich weiter stets der eine Funktionswert, etwa $\psi(x_0 x_2)$, als Limes rational-positiver Multipla $\left(\frac{m}{n}\right)_\nu$ von $\psi(x_0)$ darstellen

$$\lim_{\nu=\infty} \left(\frac{m}{n}\right)_\nu \psi(x_0) = \psi(x_0 x_2);$$

bei der Anwendung von (5) auf diese rationalen Zahlen $\left(\frac{m}{n}\right)_\nu$ erhält man daher

$$\lim_{\nu=\infty} \psi\left(x_0 \left(\frac{\xi_m}{\xi_n}\right)_\nu\right) = \lim_{\nu=\infty} \left(\frac{m}{n}\right)_\nu \psi(x_0) = \psi(x_0 x_2)$$

und wegen des bewiesenen monoton-einwertigen Charakters von $\psi(x)$

$$\lim_{\nu=\infty} x_0 \left(\frac{\xi_m}{\xi_n}\right)_\nu = x_0 x_2;$$

(7), angewendet auf $\left(\frac{\xi_m}{\xi_n}\right)_\nu$, ergibt mithin durch den Grenzübergang $\nu = \infty$ wegen der vorausgesetzten Stetigkeit von $\psi(x)$ allgemein

$$(6) \qquad \psi(x_0)\,\psi(x_0 x_1 x_2) = \psi(x_0 x_1)\,\psi(x_0 x_2).$$

Nach dem Hilfssatz 1 dieses Paragraphen mit $f(x) = \psi(x)$, $c = \psi(x_0)$, $t = x_0$, dessen Voraussetzungen hier insbesondere wegen (6) sämtlich erfüllt sind, ist daher

$$\psi(x) = \psi(x_0)\left(\frac{x}{x_0}\right)^\gamma = k x^\gamma,$$

wo γ beliebig reell*) und wegen (3) also auch

$$k = \frac{\psi(x_0)}{x_0^{\,\gamma}} \neq 0$$

beliebig reell ist.

§ 4.

Die Methode der kleinsten Quadrate auf Grund des Axiomensystems.

I. Ist die beliebige Annäherung

$$a \sim \alpha \; [p]$$

gegeben, so ist bei beliebigem reellen λ das Gewicht p' der hieraus auf Grund des Zuordnungsaxioms gewonnenen Annäherung

$$\lambda a \sim \lambda \alpha \; [p']$$

*) Man erkennt unmittelbar, daß $\gamma = 0$ dem zu beweisenden Satze nicht Genüge leistet.

nach dem Axiom der Unabhängigkeit vom Maßstabe

$$p' = pf(\lambda),$$

wo $f(\lambda)$ eine positive, eindeutige und stetige Funktion von λ ist. Ist weiter auch μ eine beliebige reelle Zahl, so ist das Gewicht p'' der Annäherung

$$\mu \cdot \lambda a \sim \mu \cdot \lambda \alpha \; [p'']$$

in analoger Weise

$$p'' = p' f(\mu) = pf(\lambda) f(\mu);$$

andererseits ist dasselbe Gewicht p'' derselben Annäherung

$$\lambda\mu \cdot a \sim \lambda\mu \cdot \alpha \; [p'']$$

nach dem Vorangehenden auch

$$p'' = pf(\lambda\mu),$$

und da nach dem Gewichtsaxiom jeder Annäherung nur ein Gewicht > 0 zugeordnet wird, so gilt bei beliebigen λ und μ

$$pf(\lambda\mu) = pf(\lambda) f(\mu)$$

oder

$$f(\lambda\mu) = f(\lambda) f(\mu);$$

da nach dem Gewichtsaxiom $f(\lambda)$ nicht verschwindet, so ist nach dem Hilfssatz 1 des § 3

$$f(\lambda) = \lambda^{\varkappa};$$

nach dem Axiom der Unabhängigkeit vom Maßstabe ist $f(\lambda) \not\equiv 1$, daher die Konstante

$$\varkappa \neq 0.$$

II. Seien irgend zwei unabhängige Annäherungen

$$a \sim \alpha \; [p], \quad b \sim \beta \; [q]$$

und deren „komponierte“ Annäherung

$$a + b \sim \alpha + \beta \; [r]$$

gegeben, so ist nach dem Axiom der Unabhängigkeit von Annäherungen das Gewicht r eindeutig bestimmt durch die Funktionalgleichung

$$\psi(r) = \psi(p) + \psi(q).$$

Nach I gilt weiter

$$\lambda a \sim \lambda \alpha \; [\lambda^{\varkappa} p], \quad \lambda b \sim \lambda \beta \; [\lambda^{\varkappa} q], \quad \lambda(a + b) \sim \lambda(\alpha + \beta) \; [\lambda^{\varkappa} r],$$

und es ist hier analog wie eben

$$\psi(\lambda^{\varkappa} r) = \psi(\lambda^{\varkappa} p) + \psi(\lambda^{\varkappa} q).$$

Wegen $\varkappa \neq 0$ durchläuft nun $\varrho = \lambda^{\varkappa}$ zugleich mit λ alle positiven Werte, und da $\psi(x)$ nach dem Axiom der Unabhängigkeit von Annäherungen eine reelle, eindeutige und stetige Funktion ist, so folgt aus dem Hilfssatze 2 des § 3

$$\psi(x)=kx^{\gamma}, \quad k\neq 0, \quad \gamma\neq 0.$$

III. Seien ferner

$$a\sim\alpha_1[1], \quad a\sim\alpha_2[1]$$

zwei unabhängige Annäherungen für dieselbe feste Größe a mit dem Gewicht 1, so folgt nach I

$$\frac{1}{2}a\sim\frac{1}{2}\alpha_1\left[\left(\frac{1}{2}\right)^{\varkappa}\right], \quad \frac{1}{2}a\sim\frac{1}{2}\alpha_2\left[\left(\frac{1}{2}\right)^{\varkappa}\right]$$

und weiter nach dem Axiom der Unabhängigkeit von Annäherungen

$$a\sim\frac{1}{2}\alpha_1+\frac{1}{2}\alpha_2[r],$$

wo für das Gewicht r nach II

$$kr^{\gamma}=k\left(\left(\frac{1}{2}\right)^{\varkappa}\right)^{\gamma}+k\left(\left(\frac{1}{2}\right)^{\varkappa}\right)^{\gamma}=2k\left(\left(\frac{1}{2}\right)^{\varkappa}\right)^{\gamma}$$

gilt. Nach dem Axiom des arithmetischen Mittels ist nun

$$r=2,$$

also folgt

$$k\,2^{\gamma}=2k\left(\left(\frac{1}{2}\right)^{\varkappa}\right)^{\gamma}$$

oder da $k\neq 0$ ist,

$$2^{\gamma}=2^{1-\varkappa\gamma};$$

hieraus folgt

$$\gamma=1-\varkappa\gamma, \quad (\varkappa+1)\gamma=1.$$

IV. Sind weiter

$$a\sim\alpha_1[1], \quad a\sim\alpha_2[1],\cdots, \quad a\sim\alpha_n[1]$$

irgend n unabhängige Annäherungen für dieselbe feste Größe a je mit dem Gewichte 1, so ergibt sich für das Gewicht r der „komponierten" Annäherung

$$na\sim\alpha_1+\alpha_2+\cdots+\alpha_n[r]$$

nach II

$$kr^{\gamma}=k\cdot 1^{\gamma}+k\cdot 1^{\gamma}+\cdots+k\cdot 1^{\gamma}=kn$$

oder wegen $k\neq 0$, $\gamma\neq 0$

$$r=n^{\frac{1}{\gamma}}.$$

Dieses Gewicht r besitzt nun nach dem Axiom der Größenordnung der Funktion $\psi(x)$ die Größenordnung $\frac{1}{n}$; also ist

$$\frac{1}{\gamma} = -1, \quad \gamma = -1$$

und nach III daher

$$\varkappa + 1 = \frac{1}{\gamma} = -1, \quad \varkappa = -2.$$

V. Seien endlich für n lineare Ausdrücke in irgend m festen Größen $a_1, a_2, \ldots, a_m$ irgend n unabhängige Annäherungen gegeben:

$$\begin{array}{c}
A_{11} a_1 + A_{21} a_2 + \cdots + A_{m1} a_m \sim l_1 \ [p_1], \\
A_{12} a_1 + A_{22} a_2 + \cdots + A_{m2} a_m \sim l_2 \ [p_2], \\
\vdots \\
A_{1n} a_1 + A_{2n} a_2 + \cdots + A_{mn} a_m \sim l_n \ [p_n],
\end{array}$$

wo $n > m$ und die A_{ik} bekannte Koeffizienten seien, so folgt hieraus, indem man diese Annäherungen mit $\lambda_1, \lambda_2, \ldots, \lambda_n$ multipliziert,

$$\begin{array}{c}
\lambda_1 (A_{11} a_1 + A_{21} a_2 + \cdots + A_{m1} a_m) \sim \lambda_1 l_1 \left[\frac{p_1}{\lambda_1^2}\right], \\
\lambda_2 (A_{12} a_1 + A_{22} a_2 + \cdots + A_{m2} a_m) \sim \lambda_2 l_2 \left[\frac{p_2}{\lambda_2^2}\right], \\
\vdots \\
\lambda_n (A_{1n} a_1 + A_{2n} a_2 + \cdots + A_{mn} a_m) \sim \lambda_n l_n \left[\frac{p_n}{\lambda_n^2}\right]
\end{array}$$

und weiter durch „Komposition“ dieser Annäherungen

$$a_1 \sum_{\nu=1}^{n} \lambda_\nu A_{1\nu} + a_2 \sum_{\nu=1}^{n} \lambda_\nu A_{2\nu} + \cdots + a_m \sum_{\nu=1}^{n} \lambda_\nu A_{m\nu} \sim \lambda_1 l_1 + \lambda_2 l_2 + \cdots + \lambda_n l_n [r],$$

wo für das Gewicht r dieser „komponierten“ Annäherung nach II und IV

$$\frac{k}{r} = \frac{k \lambda_1^2}{p_1} + \frac{k \lambda_2^2}{p_2} + \cdots + \frac{k \lambda_n^2}{p_n}$$

gilt. Sucht man nun die „beste Annäherung“ für irgendeine der m festen Größen, etwa a_μ, so muß nach dem Ausgleichungsaxiom die Komposition der gegebenen n Annäherungen mit $\lambda_1, \lambda_2, \ldots, \lambda_n$ erstens auf der linken Seite identisch a_μ liefern; es entstehen also durch Koeffizientenvergleichung von $a_1, a_2, \ldots, a_m$ die folgenden m linearen Gleichungen für $\lambda_1, \lambda_2, \ldots, \lambda_n$

$$\text{(A)} \qquad \sum_{\nu=1}^{n} \lambda_\nu A_{k\nu} = \begin{cases} 1 \text{ für } k = \mu, \\ 0 \text{ für } k = 1, 2, \cdots, m \text{ außer } k = \mu. \end{cases}$$

Zweitens besitzt die „beste Annäherung" für a_μ auch das größte Gewicht r, es sind also $\lambda_1, \lambda_2, \cdots, \lambda_n$ so zu bestimmen, daß

$$\text{(B)} \qquad \frac{\lambda_1^2}{p_1} + \frac{\lambda_2^2}{p_2} + \cdots + \frac{\lambda_n^2}{p_n} = \text{Min.}$$

wird. Auf diesen Relationen (A) und (B) beruht aber die Behandlung der Methode der kleinsten Quadrate in ihrer üblichen*) Form.

Auch der Fall der bedingten Ausgleichung, wo außer den bisherigen Annäherungen noch einige exakte lineare Beziehungen durch die Ausgleichungswerte für $a_1, a_2, \ldots, a_m$ unter allen Umständen zu erfüllen sind, läßt sich in bekannter Weise hieran anschließen.

Göttingen, den 3. März 1914.

*) Vgl. z. B. A. A. Markoff, Wahrscheinlichkeitsrechnung, deutsch von H. Liebmann, Leipzig-Berlin 1912, S. 222 ff.