

Wahrscheinlichkeit und Regression

Vorlesung WS 19/20

Multiple Lineare Regression

Prof. Dr. Rolf Steyer

Multiple Lineare Regression

Worum geht es in dieser Sitzung?

- Das Konzept der multiplen linearen Regression
- Spezialfälle
- Eigenschaften des Residuums
- Darstellung in Matrixnotation
- Identifikation der Regressionskoeffizienten
- Der multiple Determinationskoeffizient
- Multiple lineare Quasi-Regression
- Statistische Modelle zur multiplen linearen Regression
- Das Allgemeine Lineare Modell

Multiple lineare Regression: Definition

Definition 1 Seien Y und X_1, \dots, X_m numerische Zufallsvariablen auf demselben Wahrscheinlichkeitsraum mit endlichen Erwartungswerten, positiven, endlichen Varianzen, sowie invertierbarer Kovarianzmatrix $\Sigma_{\mathbf{X}\mathbf{X}}$. Dann heißt die bedingte Erwartung $E(Y | X_1, \dots, X_m)$ linear in (X_1, \dots, X_m) parametrisierbar, falls es reelle Zahlen $\beta_0, \beta_1, \dots, \beta_m$ gibt, für die gilt

$$E(Y | X_1, \dots, X_m) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m.$$

Die Funktion $g: \mathbb{R}^m \rightarrow \mathbb{R}$ mit

$$g(x_1, \dots, x_m) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m, \quad \forall (x_1, \dots, x_m) \in \mathbb{R}^m,$$

nennen wir die multiple lineare Regression von Y auf (X_1, \dots, X_m) .

Multiple lineare Regression: Spezialfälle I

Bei der obigen Definition setzen wir nicht voraus, dass die Regressoren X_1, \dots, X_m unabhängig voneinander definiert sind. In Kapitel 9 haben wir bereits darauf hingewiesen, dass z. B. die bedingte Erwartung $E(Y | X)$ linear in (X, X^2) parametrisierbar ist, falls es reelle Zahlen $\beta_0, \beta_1, \beta_2$ gibt, für die gilt:

$$E(Y | X) = \beta_0 + \beta_1 X + \beta_2 X^2.$$

Multiple lineare Regression: Spezialfälle II

Ein weiterer Spezialfall, den wir bereits im Kapitel über die bedingte lineare Regression kennen gelernt haben, ist

$$E(Y | X_1, X_2) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 \cdot X_2.$$

Auch hier ist die bedingte Erwartung $E(Y | X_1, X_2)$ linear in $(X_1, X_2, X_1 \cdot X_2)$ parametrisierbar und es gilt:

$$E(Y | X_1, X_2) = E(Y | X_1, X_2, X_1 \cdot X_2).$$

Multiple lineare Regression: Spezialfälle III

Schließlich sei explizit auch noch einmal auf den Fall eines qualitativen Regressors X mit $n + 1$ Werten x_0, x_1, \dots, x_n hingewiesen. Gilt

$$E(Y | X) = \beta_0 + \beta_1 \cdot I_1 + \beta_2 \cdot I_2 + \dots + \beta_n \cdot I_n,$$

wobei I_1, \dots, I_n die Indikatorvariablen für die Werte x_1, \dots, x_n sind, dann liegt ebenfalls ein Spezialfall der multiplen linearen Regression vor. Die bedingte Erwartung $E(Y | X)$ ist stets linear in (I_1, \dots, I_n) parametrisierbar, auch dann, wenn $E(Y | X)$ nicht linear in (X) parametrisierbar ist.

Eigenschaften des Residuums

Das Residuum

$$\varepsilon := Y - E(Y | X_1, \dots, X_m)$$

besitzt die bekannten Eigenschaften

$$\begin{aligned} E(\varepsilon) &= 0, \\ E(\varepsilon | X_1, \dots, X_m) &= 0, \\ \text{Cov}[\varepsilon, f(X_1, \dots, X_m)] &= 0, \\ \text{Cov}(\varepsilon, X_i) &= 0 \quad \text{für } i = 1, \dots, m, \end{aligned}$$

wobei $f(X_1, \dots, X_m)$ eine beliebige numerische Funktion der Regressoren bezeichnet.

Darstellung in Matrixnotation I

Fasst man die Regressoren zu dem Zeilenvektor $\mathbf{x}' = (X_1, \dots, X_m)$ und die Regressionskoeffizienten β_1, \dots, β_m zu einem m -dimensionalen Spaltenvektor $\boldsymbol{\beta} = (\beta_1, \dots, \beta_m)'$ zusammen, so können wir in Matrix- bzw. Vektornotation auch folgendermaßen schreiben:

$$E(\mathbf{y}|\mathbf{x}) = \beta_0 + \mathbf{x}'\boldsymbol{\beta} = \beta_0 + (X_1, \dots, X_m) \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}.$$

Auch hier wird der Regressand als Vektor $\mathbf{y} = (Y)$ aufgefasst, der eben nur aus einer einzigen Komponente, nämlich Y , besteht. Daher ist auch β_0 eine reelle Zahl.

Darstellung in Matrixnotation II

Definieren wir den Zeilenvektor $\mathbf{z}' := (1, X_1, \dots, X_m)$ und den Spaltenvektor $\boldsymbol{\gamma} := (\beta_0, \beta_1, \dots, \beta_m)'$, so können wir die letzte Gleichung auch wie folgt schreiben:

$$E(\mathbf{y}|\mathbf{x}) = E(\mathbf{y}|\mathbf{z}) = \mathbf{z}'\boldsymbol{\gamma} = (1, X_1, \dots, X_m) \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}.$$

Identifikation der Regressionskoeffizienten I

Zur Identifikation von β_0 und der Komponenten von $\boldsymbol{\beta} = (\beta_1, \dots, \beta_m)'$ greift man auf die Erwartungswerte des Regressanden und der Regressoren sowie die Kovarianzmatrizen $\boldsymbol{\Sigma}_{xx}$ und $\boldsymbol{\Sigma}_{xy}$ zurück. Für die Konstante β_0 ergibt sich:

$$\beta_0 = E(\mathbf{y}) - E(\mathbf{x})'\boldsymbol{\beta} = E(Y) - [E(X_1), \dots, E(X_m)] \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}.$$

Identifikation der Regressionskoeffizienten II

Multipliziert man beide Seiten der folgenden Gleichung

$$\begin{aligned} \boldsymbol{\Sigma}_{xy} &= \text{Cov}(\mathbf{x}, \mathbf{y}) = \text{Cov}(\mathbf{x}, \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \varepsilon) = \text{Cov}(\mathbf{x}, \beta_0 + \boldsymbol{\beta}'\mathbf{x} + \varepsilon) \\ &= \text{Cov}(\mathbf{x}, \mathbf{x})\boldsymbol{\beta} = \boldsymbol{\Sigma}_{xx}\boldsymbol{\beta} \end{aligned}$$

mit $\boldsymbol{\Sigma}_{xx}^{-1}$ vor, so folgt

$$\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\Sigma}_{xx}\boldsymbol{\beta} = \boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\Sigma}_{xy}.$$

Da $\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\Sigma}_{xx} = \mathbf{I}$ die Einheitsmatrix ist, folgt daraus

$$\boldsymbol{\beta} = \boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\Sigma}_{xy}$$

Damit haben wir eine Formel zur Berechnung der Regressionskoeffizienten aus den Varianzen und Kovarianzen der Regressoren und des Regressanden.

Identifikation der Regressionskoeffizienten bei zwei Regressoren I

Im Fall mit zwei Regressoren X_1 und X_2 erhält man die bereits aus Kapitel 9 bekannte Gleichung

$$\begin{aligned}\beta_0 &= E(Y) - E(\mathbf{x})'\boldsymbol{\beta} = E(Y) - [\beta_1 E(X_1) + \beta_2 E(X_2)] \\ &= E(Y) - \beta_1 E(X_1) - \beta_2 E(X_2).\end{aligned}$$

Identifikation der Regressionskoeffizienten bei zwei Regressoren II

Die Varianz-Kovarianzmatrix $\boldsymbol{\Sigma}_{xx}$ und die Kovarianzmatrix $\boldsymbol{\Sigma}_{xy}$ sind:

$$\boldsymbol{\Sigma}_{xx} = \begin{pmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) \\ \text{Cov}(X_1, X_2) & \text{Var}(X_2) \end{pmatrix} \text{ und } \boldsymbol{\Sigma}_{xy} = \begin{pmatrix} \text{Cov}(X_1, Y) \\ \text{Cov}(X_2, Y) \end{pmatrix}$$

Die Kramersche Regel aus dem letzten Kapitel liefert:

$$\boldsymbol{\Sigma}_{xx}^{-1} = \frac{1}{\text{Var}(X_1)\text{Var}(X_2) - \text{Cov}(X_1, X_2)^2} \begin{pmatrix} \text{Var}(X_2) & -\text{Cov}(X_1, X_2) \\ -\text{Cov}(X_1, X_2) & \text{Var}(X_1) \end{pmatrix}$$

Bildet man das Produkt $\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\Sigma}_{xy}$ folgen:

$$\beta_1 = \frac{\text{Var}(X_2)\text{Cov}(X_1, Y) - \text{Cov}(X_2, Y)\text{Cov}(X_1, X_2)}{\text{Var}(X_1)\text{Var}(X_2) - \text{Cov}(X_1, X_2)^2},$$

$$\beta_2 = \frac{\text{Var}(X_1)\text{Cov}(X_2, Y) - \text{Cov}(X_1, Y)\text{Cov}(X_1, X_2)}{\text{Var}(X_1)\text{Var}(X_2) - \text{Cov}(X_1, X_2)^2}.$$

Der multiple Determinationskoeffizient

Für die Varianz $Var[E(\mathbf{y}|\mathbf{x})]$ der bedingten Erwartung gilt:

$$\begin{aligned} Var[E(\mathbf{y}|\mathbf{x})] &= Var(\beta_0 + \mathbf{x}'\boldsymbol{\beta}) = Var(\mathbf{x}'\boldsymbol{\beta}) = Var(\boldsymbol{\beta}'\mathbf{x}) = \boldsymbol{\beta}'Var(\mathbf{x})\boldsymbol{\beta} = \\ &= \boldsymbol{\beta}'\boldsymbol{\Sigma}_{xx}\boldsymbol{\beta}, \end{aligned}$$

Dabei sind $Var(\mathbf{x}) = \boldsymbol{\Sigma}_{xx}$ die $(m \times m)$ Varianz-Kovarianzmatrix der Regressoren X_1, \dots, X_m und $\boldsymbol{\beta}$ der m -dimensionale Spaltenvektor der Regressionskoeffizienten β_1, \dots, β_m . Der multiple Determinationskoeffizient ergibt sich dann wie folgt:

$$R_{Y|X_1, \dots, X_m}^2 = \frac{\boldsymbol{\beta}'\boldsymbol{\Sigma}_{xx}\boldsymbol{\beta}}{Var(Y)}.$$

Der multiple Determinationskoeffizient: Spezialfälle

Im Fall mit zwei Regressoren X_1 und X_2 erhält man die bereits aus Kapitel 9 bekannte Gleichung

$$R_{Y|X_1, X_2}^2 = \frac{\beta_1^2 Var(X_1) + \beta_2^2 Var(X_2) + 2\beta_1\beta_2 Cov(X_1, X_2)}{Var(Y)}.$$

Für den Spezialfall, dass alle Regressoren *paarweise unkorreliert* sind, vereinfacht sich der Ausdruck für den multiplen Determinationskoeffizienten zu

$$R_{Y|X_1, \dots, X_m}^2 = \left(\sum_{i=1}^m \beta_i^2 Var(X_i) \right) \frac{1}{Var(Y)}.$$

Multiple lineare Quasi-Regression: Definition

Definition 2 Unter den gleichen Voraussetzungen wie in Definition 14.1 definieren wir die multiple lineare Quasi-Regression als diejenige Funktion $f: \mathbb{R}^m \rightarrow \mathbb{R}$, für die gilt

$$f(x_1, \dots, x_m) = \beta_0 + \beta_1 \cdot x_1 + \dots + \beta_m \cdot x_m, \quad \forall (x_1, \dots, x_m) \in \mathbb{R}^m,$$

sowie

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_m X_m + \nu$$

mit

$$E(\nu) = 0$$

und

$$\text{Cov}(\nu, X_1) = \dots = \text{Cov}(\nu, X_m) = 0.$$

Die Komposition von (X_1, \dots, X_m) und f notieren wir mit $Q_{lin}(Y | X_1, \dots, X_m)$ oder $Q_{lin}(\mathbf{y} | \mathbf{x})$. Es gilt dann:

$$Q_{lin}(Y | X_1, \dots, X_m) = \beta_0 + \beta_1 X_1 + \dots + \beta_m X_m.$$

Multiple lineare Quasi-Regression: Alternative Definition

Definition 3 Unter den gleichen Voraussetzungen wie zuvor, können wir $Q_{lin}(\mathbf{y} | \mathbf{x})$ auch als diejenige Linearkombination $\beta_0 + \mathbf{x}'\boldsymbol{\beta}$ definieren, welche die folgende Funktion von b_0 und \mathbf{b} , den mean-squared-error, minimiert:

$$MSE(b_0, \mathbf{b}) = E[[Y - (b_0 + \mathbf{x}'\mathbf{b})]^2].$$

Diejenige Zahl b_0 und derjenige Vektor \mathbf{b} , für welche die Funktion $MSE(b_0, \mathbf{b})$ ein Minimum annimmt, seien mit β_0 und $\boldsymbol{\beta}$, respektive, bezeichnet. Es gilt dann:

$$Q_{lin}(\mathbf{y} | \mathbf{x}) = \beta_0 + \mathbf{x}'\boldsymbol{\beta}.$$

Der Determinationskoeffizient der multiplen linearen Quasi-Regression

Für die Bestimmung der Regressionskoeffizienten der multiplen linearen Quasi-Regression gelten übrigens analog die gleichen Formeln wie für die entsprechenden Koeffizienten der echten multiplen linearen Regression. Ebenfalls gleich ist die Berechnungsformel für den Determinationskoeffizienten der linearen Quasi-Regression, d. h. es gilt

$$Q_{Y|X_1, \dots, X_m}^2 := \frac{\text{Var}[Q_{lin}(\mathbf{y}|\mathbf{x})]}{\text{Var}(Y)} = \frac{\boldsymbol{\beta}' \boldsymbol{\Sigma}_{xx} \boldsymbol{\beta}}{\text{Var}(Y)}.$$

Statistische Modelle zur multiplen linearen Regression

Bisher haben wir nur ein Einzelexperiment betrachtet: Ziehen einer Beobachtungseinheit u aus der Population und Registrierung der Werte des Regressanden und der Regressoren. Statistische Modelle beziehen sich jedoch auf N Zufallsexperimente, in denen Informationen über die zu schätzenden Parameter gesammelt werden.

Modelle mit stochastischen Regressoren

Modelle mit *stochastischen Regressoren* bestehen aus der N -maligen Wiederholung unseres bisher betrachteten Einzelexperiments. Dies führt dazu, dass man nicht mehr nur einen einzigen Regressanden Y und m Regressoren betrachten muss, sondern N Vektoren $(Y_i X_{i1} \dots X_{im})$, $i = 1, \dots, N$, die jeweils das Ergebnis des i -ten Zufallsexperiments repräsentieren. Über diese Vektoren kann man unterschiedliche Verteilungsannahmen machen, z. B. dass die $(Y_i X_{i1} \dots X_{im})$ unabhängig sind und jeder dieser Vektoren $(m + 1)$ -*variats normalverteilt* ist.

Modelle mit festen Regressoren

Mit einem anderen, weitaus häufiger verwendeten statistischen Modell, schätzt man innerhalb der Wertekombinationen x_1, \dots, x_m der Regressoren X_1, \dots, X_m die Erwartungswerte $E(Y | X_1 = x_1, \dots, X_m = x_m)$ von Y , indem man innerhalb dieser Wertekombinationen den Regressanden Y mehrfach beobachtet. Die Werte x_1, \dots, x_m der Regressoren sind dabei also nicht mehr zufällig, sondern werden als feste Größen betrachtet, die das Design des Experiments charakterisieren. Man spricht daher auch von Modellen mit festen oder nichtstochastischen Regressoren. Das wichtigste dieser Modelle mit festen Regressoren ist das *Allgemeine Lineare Modell*.

Das Allgemeine Lineare Modell

Das Allgemeine Lineare Modell (ALM) ist durch die folgenden Annahmen definiert:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}),$$

Dabei bezeichnet $\mathbf{y} = (Y_1 \dots Y_i \dots Y_N)'$ nun den Spaltenvektor der für eine Stichprobe des Gesamtumfangs N zu erhebenden "abhängigen" Variablen. Die sogenannte *Designmatrix* \mathbf{X} besteht aus $N \times (m + 1)$ festen Zahlen. Dabei besteht jede Zeile von \mathbf{X} aus den Vektoren $\mathbf{x}'_i := (1 \ x_{i1} \dots x_{im})$, eben den Wertekombinationen der Regressoren, innerhalb derer die Beobachtung Y_i erhoben wird und der vorangestellten Konstanten 1, die dazu führt, dass die Regressionskonstante β_0 die erste Komponente von $\boldsymbol{\beta} = (\beta_0 \ \beta_1 \dots \beta_m)'$ ist. Der Vektor $\boldsymbol{\varepsilon} = (\varepsilon_1 \dots \varepsilon_i \dots \varepsilon_N)'$ schließlich enthält die Residuen $\varepsilon_i := Y_i - (\mathbf{x}'_i \boldsymbol{\beta})$.

Das Allgemeine Lineare Modell II

Die zweite Annahme besagt, dass $\boldsymbol{\varepsilon}$ mit Erwartungswertvektor $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ und der $N \times N$ -Kovarianzmatrix $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}} = \sigma^2 \mathbf{I}$ multivariat normalverteilt ist. Die Residuen ε_i sind also unkorreliert und haben gleiche Varianzen. Letzteres ist die so genannte Homoskedastizitätsannahme.

Folgerungen aus den Annahmen des ALM sind zunächst:

$$E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta},$$

und

$$\boldsymbol{\Sigma}_{\mathbf{y}\mathbf{y}} = \sigma^2 \mathbf{I}.$$

Das Allgemeine Lineare Modell: Identifikationen

Ist $\mathbf{X}'\mathbf{X}$ invertierbar, so gilt

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$$

zur Schätzung des Vektors der Regressionskoeffizienten. Diese Formel erhält man durch die Minimierung der mean squared error Funktion

$$MSE(\mathbf{b}) = (\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} - \mathbf{X}\mathbf{b}).$$

Weiter ist noch

$$\Sigma_{\hat{\boldsymbol{\beta}}\hat{\boldsymbol{\beta}}} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$$

von Bedeutung, die Kovarianzmatrix dieser Schätzer. Die Wurzeln aus den diagonalen Komponenten von $\Sigma_{\hat{\boldsymbol{\beta}}\hat{\boldsymbol{\beta}}}$ sind die Standardschätzfehler der Regressionskoeffizienten.

Schätzer des Determinationskoeffizienten

Schließlich sei noch die Formel zur Schätzung des Determinationskoeffizienten genannt:

$$\hat{R}^2 = \frac{\mathbf{y}'\mathbf{X}\hat{\boldsymbol{\beta}} - N \cdot \bar{Y}^2}{\mathbf{y}'\mathbf{y} - N \cdot \bar{Y}^2} = \frac{\text{Quadratsumme der Regression}}{\text{Quadratsumme Gesamt}},$$

wobei $\bar{Y} = \frac{1}{N} \cdot \sum_{i=1}^N Y_i.$

Formulierung von Hypothesen I

Wie wir in diesem und den vorangegangenen Kapiteln gesehen haben, lassen sich mit der multiplen linearen Regression durchaus auch komplexe und nichtlineare Abhängigkeiten beschreiben. Dabei gibt es zwei grundsätzliche Strategien.

Erste Strategie: Vergleich zwischen

$$Q_{lin}(Y | X_1, \dots, X_m) = \gamma_0 + \gamma_1 X_1 + \dots + \gamma_{m-p} X_{m-p}$$

und

$$E(Y | X_1, \dots, X_m) = \beta_0 + \beta_1 X_1 + \dots + \beta_m X_m.$$

Die zu testende Nullhypothese kann dann in drei Versionen formuliert werden:

$$H_0: \beta_{m-p+1} = \beta_{m-p+2} = \dots = \beta_m = 0 \quad \text{Version 1}$$

$$H_0: Q_{lin}(Y | X_1, \dots, X_m) = E(Y | X_1, \dots, X_m) \quad \text{Version 2}$$

$$H_0: R_{Y|X_1, \dots, X_m}^2 - Q_{Y|X_1, \dots, X_m}^2 = 0 \quad \text{Version 3}$$

Formulierung von Hypothesen II

Die zweite Strategie ist noch allgemeiner: Man formuliert eine H_0 in der Form

$$H_0: \mathbf{A}\boldsymbol{\beta} - \boldsymbol{\delta} = \mathbf{0},$$

der *Allgemeinen Linearen Hypothese* und testet diese mit einem Programm wie z. B. Systat oder SPSS (über Syntax) direkt, indem man die Matrix \mathbf{A} und den Vektor $\boldsymbol{\delta}$ gemäß seiner Hypothese spezifiziert. Die Matrix \mathbf{A} muss $p \leq m$ linear unabhängige Zeilen enthalten.

Signifikanztests im ALM I

Will man die Nullhypothese $H_0: \beta_{m-p+1} = \beta_{m-p+2} = \dots = \beta_m = 0$, testen, schätzt man zunächst den Determinationskoeffizienten \hat{R}_E^2 für die Regression und dann \hat{R}_Q^2 für die multiple lineare Quasi-Regression. Dabei beachte man, dass dies jeweils mit unterschiedlichen Designmatrizen und unterschiedlichen Regressionskoeffizienten geschieht. Mit diesen Schätzungen \hat{R}_E^2 bzw. \hat{R}_Q^2 der beiden Determinationskoeffizienten geht man dann in die Formel

$$F = \frac{(\hat{R}_E^2 - \hat{R}_Q^2)/p}{(1 - \hat{R}_E^2)/(N - m - 1)},$$

die unter den Annahmen des ALM und der Gültigkeit der Nullhypothese eine F -verteilte Teststatistik liefert, mit den Zählerfreiheitsgraden $df_1 = p$ und den Nennerfreiheitsgraden $df_2 = N - m - 1$. Dabei sind m die Anzahl der Regressoren in der multiplen linearen Regression, p die Anzahl der Parameter, die laut Nullhypothese gleich null sein sollen und N der Stichprobenumfang.

Signifikanztests im ALM II

Bei der zweiten Strategie berechnet man für die jeweilige *Allgemeine Lineare Hypothese (ALH)*

$$H_0: \mathbf{A}\boldsymbol{\beta} - \boldsymbol{\delta} = \mathbf{0},$$

die Prüfgröße

$$F = \frac{\hat{Q}_h/p}{\hat{Q}_e/(N - m - 1)}.$$

Dabei sind p die Anzahl der (linear unabhängigen) Zeilen der Matrix \mathbf{A} der ALH (und damit die Anzahl der simultan geprüften Einzelhypothesen),

$$\hat{Q}_h = (\mathbf{A}\hat{\boldsymbol{\beta}} - \boldsymbol{\delta})' [\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1} (\mathbf{A}\hat{\boldsymbol{\beta}} - \boldsymbol{\delta}) \quad \text{Hypothesenquadratsumme und}$$

$$\hat{Q}_e = \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}\hat{\boldsymbol{\beta}}. \quad \text{Fehlerquadratsumme}$$

Auch die letztgenannte Prüfgröße F ist unter den Annahmen des ALM und der Gültigkeit der Nullhypothese eine F -verteilte Teststatistik, mit den Zählerfreiheitsgraden $df_1 = p$ und den Nennerfreiheitsgraden $df_2 = N - m - 1$.