Fachhochschule Nordwestschweiz
Hochschule für Wirtschaft

Saisonbereinigung und Trend einer Zeitreihe

Eine mathematisch schlüssige Methode

Studie

Autor: Helmut Vetter
Ort, Datum: Arlesheim, 06.03.2017

Diese Arbeit wurde mit TexLive erstellt.

Saisonbereinigung und Trend einer Zeitreihe
Eine mathematisch schlüssige Methode

Autor
Vetter, Helmut
Schillerweg 2
CH-4144 Arlesheim
061 599 51 09
helmut.vetter@fhnw.ch

Auftraggeberschaft
Angeregt durch
Spengler, Philipp
Fachhochschule für Wirtschaft

Basel, März 2017

Ehrenwörtliche Erklärung

Ich versichere, dass ich die vorliegende Arbeit selbstständig und ohne Benutzung anderer als der im Literaturverzeichnis angegebenen Quellen und Hilfsmittel angefertigt habe.

Die wörtlich oder inhaltlich den im Literaturverzeichnis aufgeführten Quellen und Hilfsmitteln entnommenen Stellen sind in der Arbeit als Zitat bzw. Paraphrase kenntlich gemacht.

Diese Arbeit ist noch nicht veröffentlicht worden. Sie ist somit weder anderen Interessenten zugänglich gemacht noch einer anderen Prüfungsbehörde vorgelegt worden.

Arlesheim, 06.03.2017

Helmut Vetter

Helmut Vetter

Management Summary

Es gibt verschiedene ad-hoc-Ansätze zur Lösung des Problems. Ich stelle hier eine mathematisch exakte Lösung dar. Und führe die Lösung an einer konkreten Zeitreihe aus.

Inhaltsverzeichnis

1 Einleitung

Es bezeichne $i = 0$ bis $n - 1$ aufeinanderfolgende Jahre. In Tabelle 1 sind $i = 0$ bis 4 fünf aufeinander folgende Jahre.

Das Jahr werde in m gleichlange Abschnitte unterteilt und diese durch $j = 1$ bis m nummeriert. In Tabelle 1 sind $j = 1$ bis 4 die vier Quartale des Jahres. Den Begriff 'Quartale' verwende ich nachfolgend auch falls $m \neq 4$ ist.

Zu jedem Paar (i, j) gehört ein Messwert. Einzelne Werte dürfen dabei aber auch fehlen. Dazu soll die Indikatorfunktion $\delta(i, j)$ definiert werden, die gleich 1 ist, wenn der Wert $y(i, j)$ vorhanden ist, andernfalls ist sie gleich 0 definiert.

Die entsprechenden Zeitpunkte werden mit $x(i, j) = i + \dfrac{2j-1}{2m}$ bezeichnet.

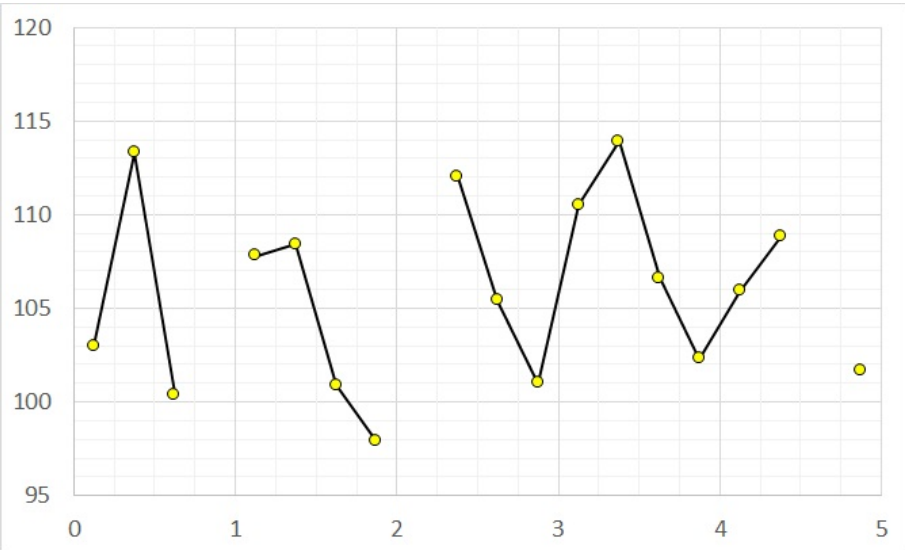
Die Messwerte werden mit $y(i, j)$ bezeichnet und sollen durch den Term $a \cdot x(i, j) + s_j$ bestmöglich (nach der Methode der kleinsten Quadrate) approximiert werden. Es sollen also a und s_j für $j = 1, \ldots, m$ so bestimmt werden, dass
$Q(a, s_1, \ldots s_m) := \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{m} \delta(i, j) \cdot (y(i, j) - (a \cdot x(i, j) + s_j))^2$ minimal wird.

Es handelt sich bei Q um eine quadratische Funktion in den $m + 1$ Variablen, a und s_1 bis s_m.

Nachfolgend soll das Problem sowohl auf analytischem als auch auf algebraischem Weg gelöst werden.

Tabelle 1 und Grafik 1: Rohdaten

i	j	$x(i,j)$	$y(i,j)$
0	1	0.125	103.0
0	2	0.375	113.3
0	3	0.625	100.4
0	4	0.875	
1	1	1.125	107.8
1	2	1.375	108.4
1	3	1.625	100.9
1	4	1.875	97.9
2	1	2.125	
2	2	2.375	112.0
2	3	2.625	105.4
2	4	2.875	101.0
3	1	3.125	110.5
3	2	3.375	113.9
3	3	3.625	106.6
3	4	3.875	102.3
4	1	4.125	105.9
4	2	4.375	108.8
4	3	4.625	
4	4	4.875	101.7



2 Analytische Lösung

Problem: Finde $m+1$ reelle Zahlen $(a,(s_j)_{j=1}^m)$, sodass $Q(a,s_1,\dots s_m) := \sum\limits_{i=1}^n \sum\limits_{j=1}^m \delta(i,j)\cdot(y(i,j)-(a\cdot x(i,j)+s_j))^2$ minimal wird.

Als notwendige Bedingung für ein lokales Minimum ergibt sich, dass die partiellen Ableitungen $\dfrac{\partial Q}{\partial a}$ und $\dfrac{\partial Q}{\partial s_j}$ für $j = 1,\dots m$ alle gleich Null sein müssen. Dies führt auf ein System von $m+1$ linearen Gleichungen.

$$\boxed{1}\ \frac{\partial}{\partial a}Q(a,s_1,\dots s_m) = \sum_{i=1}^n \sum_{j=1}^m \delta(i,j)\cdot 2\cdot(y(i,j)-(a\cdot x(i,j)+s_j))\cdot(-x(i,j)) =$$
$$= -2\sum_{i=1}^n \sum_{j=1}^m \delta(i,j)\cdot x(i,j)\cdot(y(i,j)-(a\cdot x(i,j)+s_j) \overset{!}{=} 0$$
$$\Rightarrow a\cdot\sum_{i=1}^n \sum_{j=1}^m \delta(i,j)\cdot x(i,j)^2 + \sum_{j=1}^m (s_j\cdot\sum_{i=1}^n \delta(i,j)\cdot x(i,j)) = \sum_{i=1}^n \sum_{j=1}^m \delta(i,j)\cdot x(i,j)\cdot y(i,j)$$

$$\boxed{2}\text{ bis }\boxed{\text{m+1}}\ \frac{\partial}{\partial s_j}Q(a,s_1,\dots s_m) = \sum_{i=1}^n \sum_{j=1}^m \delta(i,j)\cdot 2\cdot(y(i,j)-(a\cdot x(i,j)+s_j))\cdot(-1) =$$
$$= -2\sum_{i=1}^n \delta(i,j)\cdot(y(i,j)-(a\cdot x(i,j)+s_j) \overset{!}{=} 0 \text{ für } j = 1 \text{ bis } m$$
$$\Rightarrow a\cdot\sum_{i=1}^n \delta(i,j)\cdot x(i,j) + s_j\cdot\sum_{i=1}^n \delta(i,j) = \sum_{i=1}^n \delta(i,j)\cdot y(i,j) \text{ für } j = 1 \text{ bis } m$$

In Matrizenschreibweise ergibt sich somit folgendes lineare Gleichungssystem.

$$\begin{bmatrix} \sum\limits_{i=1}^n\sum\limits_{j=1}^m \delta(i,j)x(i,j)^2 & \sum\limits_{i=1}^n \delta(i,1)x(i,1) & \dots & \sum\limits_{i=1}^n \delta(i,m)x(i,m) \\ \sum\limits_{i=1}^n \delta(i,1)x(i,1) & \sum\limits_{i=1}^n \delta(i,1) & .0. & 0 \\ \dots & .0. & \dots & .0. \\ \sum\limits_{i=1}^n \delta(i,m)x(i,m) & 0 & .0. & \sum\limits_{i=1}^n \delta(i,m) \end{bmatrix} \cdot \begin{bmatrix} a \\ s_1 \\ \dots \\ s_m \end{bmatrix} = \begin{bmatrix} \sum\limits_{i=1}^n\sum\limits_{j=1}^m \delta(i,j)x(i,j)y(i,j) \\ \sum\limits_{i=1}^n \delta(i,1)y(i,1) \\ \dots \\ \sum\limits_{i=1}^n \delta(i,m)y(i,m) \end{bmatrix}$$

Nachfolgend wird gezeigt, dass die Koeffizientenmatrix genau dann invertierbar ist, wenn

1) für jedes 'Quartal' j mindestens ein Jahr i existiert mit $\delta(i,j) = 1$, das heisst für jedes j mindestens ein i mit einem Wert $y(i,j)$ vorliegt und

2) für mindestens ein 'Quartal' j mindestens zwei Jahre i mit $\delta(i,j) \neq 0$ existieren, das heisst für mindestens ein j mindestens zwei i mit Werten $y(i,j)$ vorliegen.

Beweis: Bezeichne $n_j := \sum\limits_{i=1}^n \delta(i,j)$ die Anzahl der vorliegenden Messungen zum 'Quartal' j.

Bezeichne $x_j := \dfrac{\sum\limits_{i=1}^n \delta(i,j)x(i,j)}{n_j}$ den Mittelwert der Zeitpunkte der vorliegenden Messungen zu 'Quartal' j.

Die Koeffizientenmatrix lässt sich damit schreiben als:

$$\begin{bmatrix} \sum\limits_{i=1}^n\sum\limits_{j=1}^m \delta(i,j)x(i,j)^2 & n_1x_1 & \dots & n_mx_m \\ n_1x_1 & n_1 & .0. & 0 \\ \dots & .0. & \dots & .0. \\ n_mx_m & 0 & .0. & n_m \end{bmatrix}$$

Bezeichne $\overrightarrow{a}$ die erste Spalte der Matrix, $\overrightarrow{b_1}$ bis $\overrightarrow{b_m}$ die restlichen m Spalten der Matrix. Falls für jedes 'Quartal' j Messwerte vorliegen, also alle $n_j > 0$ sind, so sind die Spalten $\overrightarrow{b_j}$ unabhängig. Ansonsten ist eine Spalte $\overrightarrow{b_j}$ eine Nullspalte und die Spalten also linear abhängig.

Jetzt wird vorausgesetzt, dass alle $n_j > 0$ sind. Die erste Spalte $\overrightarrow{a}$ lässt sich dann wie folgt reduzieren $\overrightarrow{a'} :=$

$\overrightarrow{a} - \sum\limits_{j=1}^{m} x_j \overrightarrow{b_j}$. $\overrightarrow{a'}$ hat in den Zeilen 2 bis $m+1$ alles Nullen. Der Eintrag der ersten Zeile ist

$\overrightarrow{a'}(1) = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} \delta(i,j)x(i,j)^2 - \sum\limits_{j=1}^{m} n_j x_j^2 = \sum\limits_{j=1}^{m} n_j \cdot (\sum\limits_{i=1}^{n} \frac{\delta(i,j)}{n_j} x(i,j)^2 - x_j^2)$.

Weil die Funktion $f(x) = x^2$ konvex ist, gilt für jedes j dass $\sum\limits_{i=1}^{n} \frac{\delta(i,j)}{n_j} x(i,j)^2 \geq (\sum\limits_{i=1}^{n} \frac{\delta(i,j)}{n_j} x(i,j))^2 = x_j^2$. Gleichheit gilt hierbei nur, wenn alle Zeitpunkte $x(i,j)$ mit $\delta(i,j) \neq 0$ denselben Wert haben, das heisst es nur ein i mit $\delta(i,j) \neq 0$ gibt.

Sind die n_1 bis n_m alle ungleich Null, so ist äquivalent:
$\overrightarrow{a}$, $\overrightarrow{b_1}$, $\ldots \overrightarrow{b_m}$ sind unabhängig $\Leftrightarrow$
$\overrightarrow{a'}$, $\overrightarrow{b_1}$, $\ldots \overrightarrow{b_m}$ sind unabhängig $\Leftrightarrow$
$\overrightarrow{a'}(1) = \sum\limits_{j=1}^{m} n_j \cdot (\sum\limits_{i=1}^{n} \frac{\delta(i,j)}{n_j} x(i,j)^2 - x_j^2) > 0 \Leftrightarrow$
Es gibt mindestens ein 'Quartal' j mit mindestens zwei Jahren i mit $\delta(i,j) \neq 0$. qed.

Sind die Voraussetzungen 1) und 2) erfüllt, ergibt sich eine eindeutige Lösung $(a, s_1, \ldots, s_m)$.

Um nachzuweisen, dass es sich hierbei um ein globales Minimum der quadratischen Funktion Q handelt, ist der algebraische Weg der geeignete! $\Rightarrow$ Algebraische Lösung.

3 Numerische Auswertung

An den Daten von Tabelle 1 ergibt sich als lineares Gleichungssystem

$$\begin{bmatrix} 139.890625 & 8.5 & 11.875 & 8.5 & 13.5 \\ 8.5 & 4 & 0 & 0 & 0 \\ 11.875 & 0 & 5 & 0 & 0 \\ 8.5 & 0 & 0 & 4 & 0 \\ 13.5 & 0 & 0 & 0 & 4 \end{bmatrix} \cdot \begin{bmatrix} a \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix} = \begin{bmatrix} 4490.2 \\ 427.2 \\ 556.4 \\ 413.3 \\ 402.9 \end{bmatrix}$$

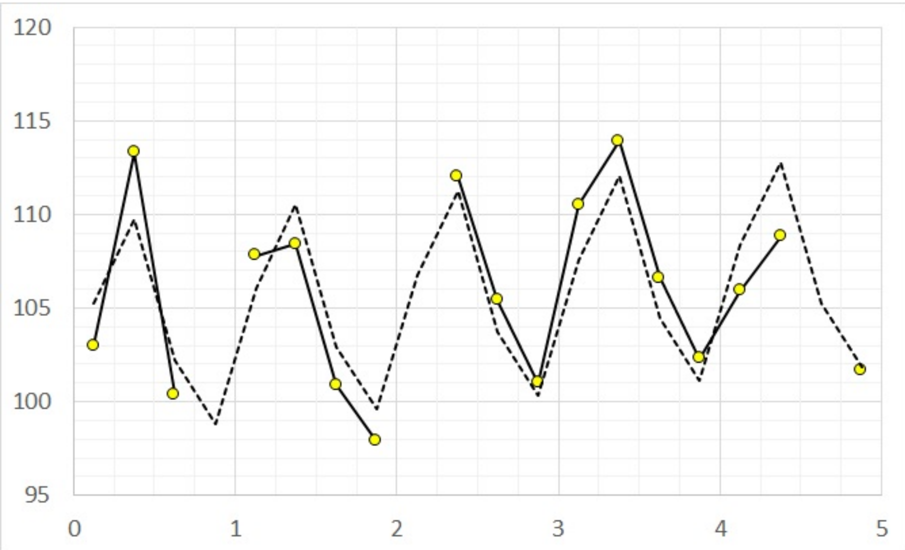
Die Koeffizientenmatrix ist invertierbar.

Als Lösung ergibt sich an den Daten von Tabelle 1:

$$\begin{bmatrix} a \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix} = \begin{bmatrix} 0.7633 \\ 105.1779 \\ 109.4671 \\ 101.7029 \\ 98.1487 \end{bmatrix}$$

Trägt man die Wertepaare $(x(i,j), y(i,j))$ bzw. $(x(i,j), a \cdot x(i,j) + s_j)$ in ein Koordinatensystem ein und verbindet die Punkte fortlaufend durch Strecken, so erhält man folgendes Bild.

Grafik 2: Rohdaten (ausgezogen) und ausgeglichene Daten (gestrichelt)



4 Algebraische Lösung

Bezeichne $J := \{(i,j) \mid y(i,j)$ liegt vor$\}$.

Algebraisch gesehen soll der Vektor $\overrightarrow{y} = [(i,j) \mapsto y(i,j)] \in \mathbb{R}^J$ bestmöglich (in der euklideschen Norm) durch eine Linearkombination der Vektoren $\overrightarrow{x} := [(i,j) \mapsto x(i,j)]$ und $\overrightarrow{\sigma_k} := [(i,j) \mapsto \chi(j=k)]$ für $1 \le k \le m$ approximiert werden.

Bezeichne $< \overrightarrow{u} \mid \overrightarrow{v} >= \overrightarrow{u}^\mathsf{T}\overrightarrow{v}$ das Standardskalarprodukt im $\mathbb{R}^J$.

Bezeichne X die Matrix bestehend aus den $m+1$ Spaltenvektoren $[\overrightarrow{x}, \overrightarrow{\sigma_1}, \overrightarrow{\sigma_2}, \ldots, \overrightarrow{\sigma_m}]$.

X definiert via $c \mapsto X \cdot c$ eine lineare Abbildung von $\mathbb{R}^{m+1}$ nach $\mathbb{R}^J$.

Behauptung 1: $\mathrm{Im}(X)^\perp = \mathrm{Ker}(X^\mathsf{T})$

Beweis: $\overrightarrow{v} \in \mathrm{Im}(X)^\perp \Leftrightarrow$ für alle $\overrightarrow{w}$ gilt $< \overrightarrow{v} \mid X\overrightarrow{w} >= \overrightarrow{v}^\mathsf{T} X \overrightarrow{w} = 0 \Leftrightarrow$ für alle $\overrightarrow{w}$ gilt $< X^\mathsf{T}\overrightarrow{v} \mid \overrightarrow{w} >= (X^\mathsf{T}\overrightarrow{v})^\mathsf{T}\overrightarrow{w} = 0 \Leftrightarrow X^\mathsf{T}\overrightarrow{v} = 0 \Leftrightarrow \overrightarrow{v} \in \mathrm{Ker}(X^\mathsf{T})$ qed.

Es gilt $\mathbb{R}^J = \mathrm{Im}(X) \oplus \mathrm{Im}(X)^\perp$ wie man leicht über das Orthogonalisierungsverfahren von Gram-Schmidt zeigt. Jedes $\overrightarrow{y}$ hat also eine eindeutige Darstellung $\overrightarrow{y} = X\overrightarrow{c} + \overrightarrow{u}$ mit $\overrightarrow{u} \in \mathrm{Im}(X)^\perp$.

Definition: $\overrightarrow{u}$ heisst das Residuum. $\overrightarrow{u}$ ist eindeutig bestimmt. $\overrightarrow{u} \in \mathrm{Im}(X)^\perp = \mathrm{Ker}(X^\mathsf{T})$

$\overrightarrow{y} = X \cdot c + \overrightarrow{u} \mid X^\mathsf{T}$
$X^\mathsf{T}\overrightarrow{y} = X^\mathsf{T}X \cdot c + \overrightarrow{0}$
$X^\mathsf{T}\overrightarrow{y} = X^\mathsf{T}X \cdot c$

Behauptung 2: X injektiv $\Leftrightarrow X^\mathsf{T}X$ ist invertierbar.
Beweis: X injektiv $\Leftrightarrow$ für $\overrightarrow{v} \neq \overrightarrow{0}$ ist $X\overrightarrow{v} \neq \overrightarrow{0} \Leftrightarrow$ für $\overrightarrow{v} \neq \overrightarrow{0}$ ist $< X\overrightarrow{v} \mid X\overrightarrow{v} >= \overrightarrow{v}^\mathsf{T}X^\mathsf{T}X\overrightarrow{v} =< \overrightarrow{v} \mid X^\mathsf{T}X\overrightarrow{v} >\neq 0 \Rightarrow$ für $\overrightarrow{v} \neq \overrightarrow{0}$ ist $X^\mathsf{T}X\overrightarrow{v} \neq \overrightarrow{0} \Leftrightarrow X^\mathsf{T}X$ ist injektiv $\Leftrightarrow X^\mathsf{T}X$ (Endomorphismus) ist bijektiv $\Rightarrow$ X ist injektiv qed.

Leicht einzusehen: Falls für jedes 'Quartal' j mindestens ein Messwert $y(i,j)$ vorliegt und es zu mindestens einem 'Quartal' j zu zwei verschiedenen Jahren i Messwerte $y(i,j)$ gibt, dann ist X injektiv.

$X^\mathsf{T}\overrightarrow{y} = X^\mathsf{T}X \cdot c \mid (X^\mathsf{T}X)^{-1}\cdot$
$\underline{\underline{\overrightarrow{c} = (X^\mathsf{T}X)^{-1} \cdot X^\mathsf{T}\overrightarrow{y}}}$

Beweis der Minimaleigenschaft von $\overrightarrow{c}$ für $|X\overrightarrow{c} - \overrightarrow{y}|^2$

Wähle einen beliebigen Vektor $\overrightarrow{d}$ anstelle von $\overrightarrow{c}$

$|X\overrightarrow{d} - \overrightarrow{y}|^2 =$
$< X\overrightarrow{d} - \overrightarrow{y} \mid X\overrightarrow{d} - \overrightarrow{y} >=$
$< X\overrightarrow{d} - X\overrightarrow{c} + X\overrightarrow{c} - \overrightarrow{y} \mid X\overrightarrow{d} - X\overrightarrow{c} + X\overrightarrow{c} - \overrightarrow{y} >=$
$< X(\overrightarrow{d} - \overrightarrow{c}) + X\overrightarrow{c} - \overrightarrow{y} \mid X(\overrightarrow{d} - \overrightarrow{c}) + X\overrightarrow{c} - \overrightarrow{y} >=$
$< X(\overrightarrow{d} - \overrightarrow{c}) \mid X(\overrightarrow{d} - \overrightarrow{c}) > +2 < X(\overrightarrow{d} - \overrightarrow{c}) \mid X\overrightarrow{c} - \overrightarrow{y} > + < X\overrightarrow{c} - \overrightarrow{y} \mid X\overrightarrow{c} - \overrightarrow{y} >=$
$|X(\overrightarrow{d} - \overrightarrow{c})|^2 + 2 < (\overrightarrow{d} - \overrightarrow{c}) \mid X^\mathsf{T}(X\overrightarrow{c} - \overrightarrow{y}) > + |X\overrightarrow{c} - \overrightarrow{y}|^2 =$
$|X(\overrightarrow{d} - \overrightarrow{c})|^2 + 2 < (\overrightarrow{d} - \overrightarrow{c}) \mid X^\mathsf{T}X\overrightarrow{c} - X^\mathsf{T}\overrightarrow{y} > + |X\overrightarrow{c} - \overrightarrow{y}|^2 =$
$|X(\overrightarrow{d} - \overrightarrow{c})|^2 + 2 < (\overrightarrow{d} - \overrightarrow{c}) \mid \overrightarrow{0}) > + |X\overrightarrow{c} - \overrightarrow{y}|^2 =$
$|X(\overrightarrow{d} - \overrightarrow{c})|^2 + |X\overrightarrow{c} - \overrightarrow{y}|^2 \geq$
$|X\overrightarrow{c} - \overrightarrow{y}|^2$ qed.

5 Numerische Auswertung

Vektoren

(i,j)	$\overrightarrow{x}$	$\overrightarrow{\sigma_1}$	$\overrightarrow{\sigma_2}$	$\overrightarrow{\sigma_3}$	$\overrightarrow{\sigma_4}$	$\overrightarrow{y}$
$(0,1)$	0.125	1	0	0	0	103.0
$(0,2)$	0.375	0	1	0	0	113.3
$(0,3)$	0.625	0	0	1	0	100.4
$(0,4)$						
$(1,1)$	1.125	1	0	0	0	107.8
$(1,2)$	1.375	0	1	0	0	108.4
$(1,3)$	1.625	0	0	1	0	100.9
$(1,4)$	1.875	0	0	0	1	97.9
$(2,1)$						
$(2,2)$	2.375	0	1	0	0	112.0
$(2,3)$	2.625	0	0	1	0	105.4
$(2,4)$	2.875	0	0	0	1	101.0
$(3,1)$	3.125	1	0	0	0	110.5
$(3,2)$	3.375	0	1	0	0	113.9
$(3,3)$	3.625	0	0	1	0	106.6
$(3,4)$	3.875	0	0	0	1	102.3
$(4,1)$	4.125	1	0	0	0	105.9
$(4,2)$	4.375	0	1	0	0	108.8
$(4,3)$						
$(4,4)$	4.875	0	0	0	1	101.7

Daraus ergeben sich

$$X = \begin{bmatrix} 0.125 & 1 & 0 & 0 & 0 \\ 0.375 & 0 & 1 & 0 & 0 \\ 0.625 & 0 & 0 & 1 & 0 \\ 1.125 & 1 & 0 & 0 & 0 \\ 1.375 & 0 & 1 & 0 & 0 \\ 1.625 & 0 & 0 & 1 & 0 \\ 1.875 & 0 & 0 & 0 & 1 \\ 2.375 & 0 & 1 & 0 & 0 \\ 2.625 & 0 & 0 & 1 & 0 \\ 2.875 & 0 & 0 & 0 & 1 \\ 3.125 & 1 & 0 & 0 & 0 \\ 3.375 & 0 & 1 & 0 & 0 \\ 3.625 & 0 & 0 & 1 & 0 \\ 3.875 & 0 & 0 & 0 & 1 \\ 4.125 & 1 & 0 & 0 & 0 \\ 4.375 & 0 & 1 & 0 & 0 \\ 4.875 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad \overrightarrow{y} = \begin{bmatrix} 103.0 \\ 113.3 \\ 100.4 \\ 107.8 \\ 108.4 \\ 100.9 \\ 97.9 \\ 112.0 \\ 105.4 \\ 101.0 \\ 110.5 \\ 113.9 \\ 106.6 \\ 102.3 \\ 105.9 \\ 108.8 \\ 101.7 \end{bmatrix}$$

Und daraus schliesslich (wie in der analytischen Lösung)

$$c = (X^{\mathsf{T}} X)^{-1} \cdot X^{\mathsf{T}} \overrightarrow{y} = \underline{\underline{\begin{bmatrix} 0.7633 \\ 105.1779 \\ 109.4671 \\ 101.7029 \\ 98.1487 \end{bmatrix}}}$$