Da R eine obere Trapezmatrix ist, hat das Produkt $R\vec{x}$ ausgeschrieben die Treppengestalt, die der GAUSSS-Algorithmus produziert, und die rechte Seite läßt sich für jede neue rechte Seite zum Preis einer einzigen Matrixmultiplikation berechnen. Im Gegensatz zur LR-Zerlegung ist also keine Matrixinversion nötig, und dazu ist diese Methode auch noch numerisch stabiler, falls man mit Gleitkommazahlen rechnet und einen guten numerischen Algorithmus zur Berechnung der QR-Zerlegung verwendet. (GRAM-SCHMIDT ist numerisch eher nicht zu empfehlen.)

Der wesentliche Grund für die guten numerischen Eigenschaften orthogonaler und unitärer Matrizen besteht darin, daß sie Längen respektieren. Um das einzusehen, beweisen wir zunächst ein allgemeines Lemma (das auch der historische Grund für die Bezeichnung „adjungierte Matrix" ist):

Lemma: Sei $k = \mathbb{R}$ oder $\mathbb{C}$ und $A \in k^{n\times m}$. Dann ist

$$(A\vec{v}) \cdot \vec{w} = \vec{v} \cdot (A^*\vec{w}) \quad \text{für alle } \vec{v} \in k^m \text{ und } \vec{w} \in k^n ,$$

wobei links das (HERMITEsche) Standardskalarprodukt von k^n steht und rechts das von k^m.

Beweis: Es genügt, den HERMITEschen Fall zu betrachten. Dazu fassen wir einen Vektor $\vec{w} \in \mathbb{C}^n$ auf als eine $n \times 1$-Matrix $w_M \in \mathbb{C}^{n\times 1}$; für einen weiteren Vektor $\vec{u} \in \mathbb{C}^n$, aufgefaßt als Matrix $u_M \in \mathbb{C}^{n\times 1}$, ist dann das HERMITEsche Skalarprodukt $\vec{v} \cdot \vec{w}$ gleich dem Matrixprodukt ${}^t v_M \cdot \overline{w_M}$.

Ganz entsprechend ordnen wir dem Vektor $\vec{v} \in \mathbb{C}^m$ eine $m \times 1$-Matrix $v_M \in \mathbb{C}^{m\times 1}$ zu; die Matrix zum Vektor $A\vec{v}$ ist dann die Produktmatrix Av_M.

Somit ist

$$\begin{aligned}(A\vec{v}) \cdot \vec{w} &= {}^t(Av_M) \cdot \overline{w_M} = {}^t v_M \cdot {}^t A \cdot \overline{w_M} = {}^t v_M \cdot \overline{\overline{{}^t A} \cdot w_M} \\ &= {}^t v_M \cdot \overline{A^* \cdot w_M} = \vec{v} \cdot (A^*\vec{w}) .\end{aligned}$$

■

Als mehr oder weniger unmittelbare Folgerung erhalten wir:

Satz: *a)* Eine Matrix $A \in \mathbb{R}^{n\times n}$ ist genau dann orthogonal, wenn für das Standardskalarprodukt des $\mathbb{R}^n$ gilt: $(A\vec{v}) \cdot (A\vec{w}) = \vec{v} \cdot \vec{w}$ für alle $\vec{v}, \vec{w} \in \mathbb{R}^n$.
b) Eine Matrix $A \in \mathbb{C}^{n\times n}$ ist genau dann unitär, wenn für das HERMITEsche Standardprodukt des $\mathbb{C}^n$ gilt: $(A\vec{v}) \cdot (A\vec{w}) = \vec{v} \cdot \vec{w}$ für alle $\vec{v}, \vec{w} \in \mathbb{C}^n$.

Beweis: Nach dem gerade bewiesenen Lemma ist

$$(A\vec{v}) \cdot (A\vec{w}) = \vec{v} \cdot \big(A^*(A\vec{w})\big) = \vec{v} \cdot \big((A^*A)\vec{w}\big) .$$

A ist genau dann orthogonal *bzw.* unitär, wenn AA^* gleich der Einheitsmatrix ist; in diesem Fall ist $(A^*A)\vec{w} = \vec{w}$ und damit $(A\vec{v})\cdot(A\vec{w}) = \vec{v}\cdot\vec{w}$.

Falls wir umgekehrt wissen, daß $(A\vec{v})\cdot(A\vec{w}) = \vec{v}\cdot\vec{w}$ ist für alle Vektoren $\vec{v}, \vec{w}$, so ist auch $\vec{v} \cdot \big((A^*A)\vec{w}\big) = \vec{v} \cdot \vec{w}$, insbesondere also

$$\vec{e}_i \cdot \big((A^*A)\vec{e}_j\big) = \vec{e}_i \cdot \vec{e}_j = \delta_{ij}$$

für die Koordinateneinheitsvektoren.

$(A^*A)\vec{e}_j$ ist die j-te Spalte der Matrix A^*A, ihr Skalarprodukt mit $\vec{e}_i$ also der ij-Eintrag von A^*A. Da dieser gleich δ_{ij} sein muß, ist also $A^*A = E$ und A somit orthogonal *bzw.* unitär.

■

h) Orthogonale Projektionen

Ist U ein r-dimensionaler Untervektorraum eines n-dimensionalen Vektorraums V, so können wie jede Basis $\{\vec{b}_1, \dots, \vec{b}_r\}$ von U ergänzen zu einer Basis $\{\vec{b}_1, \dots, \vec{b}_n\}$ von V. Der von $\vec{b}_{r+1}$ bis $\vec{b}_n$ erzeugte Untervektorraum $W \le V$ hat dann die Eigenschaft, daß $U \cup W$ der Nullraum ist, während $U \cup W$ dem gesamten Vektorraum V erzeugt. Einen solchen Untervektorraum W bezeichnen wir als *Komplement* von U; es ist natürlich, genau wie seine Basisvektoren $\vec{b}_{r+1}$ bis $\vec{b}_n$, alles andere als eindeutig bestimmt.

Für EUKLIDische und HERMITEsche Vektorräume können wir allerdings jedem Untervektorraum ein wohlbestimmtes ausgezeichnetes Komplement zuordnen, das *orthogonale Komplement.*

Definition: V sei ein EUKLIDischer oder HERMITEscher Vektorraum und $U \leq V$ sei ein Untervektorraum von V. Das orthogonale Komplement von U ist der Untervektorraum

$$U^\perp \underset{\text{def}}{=} \left\{\vec{v} \in V \;\middle|\; \vec{u} \cdot \vec{v} = 0 \text{ für alle } \vec{u} \in U\right. .$$

Wegen der Linearität des EUKLIDischen wie auch HERMITEschen Skalarprodukts im ersten Argument ist klar, daß $U^\perp$ ein Untervektorraum von V ist. Außerdem ist klar, daß es reicht die Bedingung $\cdot\vec{u} = 0$ für die Vektoren $\vec{U}$ aus einer Basis von U nachzurechnen, denn jedes Element von U ist Linearkombination von Basisvektoren, und wenn auch im HERMITEschen Fall beim zweiten Argument die Koeffizienten komplex konjugiert werden müssen, folgt doch sofort, daß dann auch das Produkt mit einer Linearkombination der Basisvektoren verschwindet.

Lemma: U sei ein Untervektorraum des n-dimensionalen EUKLIDischen oder HERMITEschen Vektorraums V, und $\{\vec{b}_1, \dots, \vec{b}_r\}$ sei eine Orthogonalbasis von U. Ergänzt man diese zu einer Orthogonalbasis $\{\vec{b}_1, \dots, \vec{b}_n\}$ von V, so ist $\{\vec{b}_{r+1}, \dots, \vec{b}_n\}$ eine Orthogonalbasis von $U^\perp$. Insbesondere hat also das orthogonale Komplement eines r-dimensionalen Untervektorraum die Dimension $n - r$.

Beweis: Nach dem gerade Gesagten liegt ein Vektor $\vec{v} = \lambda_1\vec{b}_1 + \cdots + \lambda_n\vec{b}_n$ aus V genau dann in $U^\perp$, wenn für alle $i \leq r$ gilt:

$$\vec{v} \cdot \vec{b}_i = \left(\sum_{j=1}^{n} \lambda_j \vec{b}_j\right) \cdot \vec{b}_i = \sum_{j=1}^{n} \lambda_j \vec{b}_j \cdot \vec{b}_i = \lambda_i \vec{b}_i \cdot \vec{b}_i = 0 .$$

Da $\vec{b}_i$ als Basisvektor nicht der Nullvektor sein kann, ist $\vec{b}_i \cdot \vec{b}_i \neq 0$; daher ist dies äquivalent zum Verschwinden aller λ_i mit $i \leq r$, also zur Darstellbarkeit von $\vec{v}$ als Linearkombination der Vektoren $\vec{b}_{r+1}, \dots, \vec{b}_n$. Als Teil einer Basis sind diese linear unabhängig, also Basis ihres Erzeugnisses $U^\perp$. ∎

Korollar: *a)* V sei ein endlichdimensionaler EUKLIDischer oder HERMITEscher VektorraumVektorraum und $U \leq V$ sei ein Untervektorraum.

Dann läßt sich jedes Element $\vec{v} \in V$ eindeutig schreiben als $\vec{v} = \vec{u} + \vec{w}$ mit $\vec{u} \in U$ und $\vec{w} \in U^\perp$.
b) $U^{\perp\perp} = U$

Beweis: a) Wir wählen eine Orthogonalbasis $\{\vec{b}_1, \dots, \vec{b}_r\}$ von U und ergänzen sie zu einer Orthogonalbasis $\{\vec{b}_1, \dots, \vec{b}_n\}$ von V; nach dem Lemma ist dann $\{\vec{b}_{r+1}, \dots, \vec{b}_n\}$ eine Orthogonalbasis von $U^\perp$. Schreiben wir $\vec{v} = v_1\vec{b}_1 + \cdots + v_n\vec{b}_n$, so ist also

$$\vec{u} \underset{\text{def}}{=} v_1\vec{b}_1 + \cdots + v_r\vec{b}_r \in U, \quad \vec{w} \underset{\text{def}}{=} v_{r+1}\vec{b}_{r+1} + \cdots + v_n\vec{b}_n \in U^\perp$$

und $\vec{v} = \vec{u} + \vec{w}$. Diese Darstellung ist eindeutig wegen der Eindeutigkeit der Basisdarstellungen in V, U und $U^\perp$.

b) Für $\vec{u} \in U$ und $\vec{w} \in U^\perp$ verschwindet nach Definition von $U^\perp$ das Produkt $\vec{w} \cdot \vec{u}$, also wegen dessen (HERMITEscher) Symmetrie auch $\vec{u} \cdot \vec{w}$. Damit ist

$$\vec{u} \in U^{\perp\perp} = \left\{\vec{v} \in V \;\middle|\; \vec{v} \cdot \vec{w} = 0 \text{ für alle } \vec{w} \in U^\perp\right. ,$$

also liegt U in $U^{\perp\perp}$. Nach dem Lemma ist

$$\dim U^{\perp\perp} = \dim V - \dim U^\perp = \dim V - (\dim V - \dim U) = \dim U ,$$

also muß $U = U^{\perp\perp}$ sein ∎

Bemerkung: Tatsächlich gilt dieses Korollar auch für unendlichdimensionale Vektorräume; da die Existenz von wie auch der Umgang mit Basen im Unendlichdimensionalen etwas problematisch ist, soll aber hier, wie bereits mehrfach in diesem Skriptum, der endlichdimensionale Fall genügen.

Definition: V sei ein endlichdimensionaler EUKLIDischer oder HERMITEscher VektorraumVektorraum und $U \leq V$ sei ein Untervektorraum. Die Abbildung $\pi_U : V \to U$, die jedem Vektor $\vec{v} = \vec{u} + \vec{w} \in V$ mit $\vec{u} \in U$ und $\vec{w} \in U^\perp$ den Vektor $\vec{u}$ zuordnet, heißt *orthogonale Projektion* von V nach U.

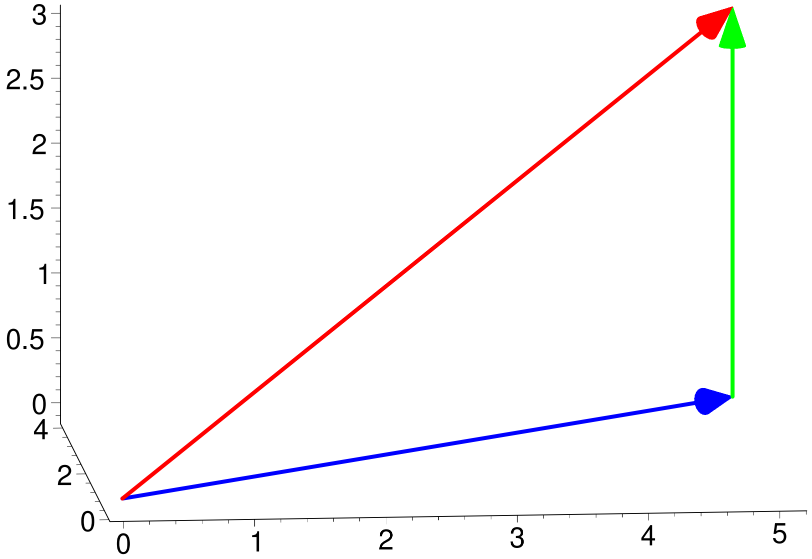

Abb. 15: Orthogonale Projektion eines Vektors

Wegen der eindeutigen Zerlegbarkeit eines Vektors in eine Komponente aus U und eine aus $U^\perp$ ist π_U offensichtlich wohldefiniert und linear; der Kern von π_U ist $U^\perp$.

Orthogonale Projektionen sind aus der Geometrie bekannt, beispielsweise als Grundriß, Aufriß und Kreuzriß eines dreidimensionalen Körpers; uns interessiert hier vor allem ihre folgende Eigenschaft:

Lemma: V sei ein endlichdimensionaler EUKLIDischer oder HERMITEscher VektorraumVektorraum, $U \le V$ ein Untervektorraum und $\vec{\in} V$. Dann ist für jeden Vektor $\vec{u} \in$

$$|\vec{v} - \vec{u}| \le |\vec{v} - \pi_U(\vec{v})| \ ,$$

d.h. $\pi_U(\vec{v})$ ist derjenige Vektor aus U, dessen Differenz mit $\vec{v}$ am kürzesten ist.

Beweis: Wir schreiben $\vec{v} = \vec{p} + \vec{w}$ mit $\vec{p} = \pi_U(\vec{v}) \in U$ und $\vec{w} \in U^\perp$. Für

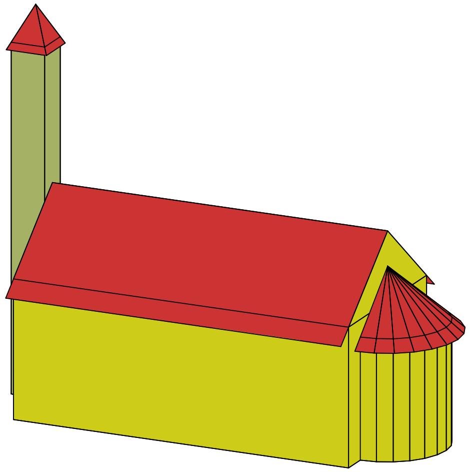

Abb. 16: Ein dreidimensionales Objekt

Abb. 17: Der Grundriß des Objekts aus Abbildung 16

jeden Vektor $\vec{u} \in U$ ist dann

$$\begin{aligned}
|\vec{v} - \vec{u}|^2 &= |\vec{p} + \vec{w} - \vec{u}|^2 = |(\vec{p} - \vec{u}) + \vec{w}|^2 \\
&= (\vec{p} - \vec{u}) \cdot (\vec{p} - \vec{u}) + \vec{w} \cdot (\vec{p} - \vec{u}) + (\vec{p} - \vec{u}) \cdot \vec{w} + \vec{w} \cdot \vec{w} \\
&= |\vec{p} - \vec{u}|^2 + |\vec{w}|^2 \ ,
\end{aligned}$$

denn $\vec{p} - \vec{u}$ liegt in U und $\vec{w}$ in $U^\perp$. Also ist $|\vec{v} - \vec{u}|$ nie kleiner als $|\vec{v} - \vec{p}|$, und die beiden Vektoren sind genau dann gleich lang, wenn

$\vec{u} - \vec{p}$ der Nullvektor ist, also $\vec{u} = \pi_U(\vec{v})$. ■

Die orthogonale Projektion auf einen Untervektorraum entspricht also geometrisch der Konstruktion des Lotfußpunkts in U.

i) Die Methode der kleinsten Quadrate

Oftmals ist zu gegebenen Beobachtungsdaten grundsätzlich bekannt, welcher *Art* von Gesetz sie genügen sollten; das Problem besteht „nur“ noch darin, die in diesem Gesetz vorkommenden *Parameter* zu bestimmen. Im einfachsten Fall könnte man etwa an einen Widerstand denken, der dadurch gemessen wird, daß man verschiedene Spannungen U_i anlegt und die zugehörigen Stromstärken I_i mißt. Nach dem Ohmschen Gesetz ist dann $U_i = R \cdot I_i$, aber aufgrund der unvermeidlichen Meßfehler werden die verschiedenen Quotienten U_i/I_i natürlich nicht alle gleich sein. Die Lösung dieses Problems ist klar: Man nimmt den Mittelwert der Quotienten. Schwieriger wird es, wenn mehrere Parameter ins Spiel kommen, wenn die Meßreihe als mehr als nur einen Parameter bestimmen soll.

Solche Fälle treten nicht nur auf in Naturwissenschaft und Technik, sondern auch in den Wirtschafts- und Sozialwissenschaften, wo es zwar selten *exakte* Gesetze gibt, man den Zusammenhang zwischen verschiedenen Größen aber trotzdem zumindest näherungsweise durch eine mathematische Formel beschreiben will – auch wenn diese in konkreten Einzelfällen gelegentlich ziemlich falsch sein kann.

Als Beispiel dieser Art können wir den Zusammenhang zwischen Korruption und Wohlstand in verschiedenen Staaten betrachten: edes Jahr veröffentlicht die Organisation *Transparancy International* ihren *corruption perceptions index (CPI),* in dem jedem Land eine Zahl zwischen null und zehn zugeordnet wird, je nachdem, wie stark Geschäftsleute, Risikospezialisten und die Bevölkerung die Korruption im betreffenden Land einschätzen: Ein Index von zehn bedeutet, daß es praktische keine Korruption gibt, während bei null nichts läuft ohne Bimbes. Die neuesten Daten stammen vom 28. August 2002 und sind unter

`http://www.transparency.org/cpi/2002/cpi2002.de.html`

zu finden. Die Zahlen werden als Mittelwerte über die letzten drei Jahren berechnet, so daß singuläre Ereignisse eines Jahres nicht zu sehr ins Gewicht fallen.

Wir vergleichen diese Zahlen mit dem Bruttoinlandsprodukt pro Einwohner, wie es die Weltbank für 2000 festgestellt hat. Dies sind die neuesten verfügbaren Zahlen; sie sind beispielsweise auf dem Server des Statistischen Bundesamtes unter

`http://www.destatis.de/ausl_prog/suche_ausland.htm`

zu finden, indem man unter „Indikatoren“ das Feld „BIP je Einwohner (real)“ auswählt. In der folgenden Tabelle sind Staaten aufgelistet, für die sowohl das Bruttosozialprodukt pro Einwohner als auch der CPI für 2002 vorliegt; das Bruttosozialprodukt pro Einwohner in US-$ ist kursiv gedruckt, der Korruptionsindex fett.

Ägypten	*1226*	**3,4**
Albanien	*899*	**2,5**
Angola	*506*	**1,7**
Argentinien	*7933*	**2,8**
Aserbaidschan	*506*	**2,0**
Äthiopien	*116*	**3,5**
Australien	*23598*	**8,6**
Bangladesch	*373*	**1,2**
Belgien	*31017*	**7,1**
Bolivien	*952*	**2,2**
Botsuana	*1526*	**6,4**
Brasilien	*4624*	**4,0**
Bulgarien	*1549*	**4,0**
Chile	*5354*	**7,5**
China	*824*	**3,5**
Costa Rica	*3912*	**4,5**
Côte d'Ivoire	*743*	**2,7**
Dänemark	*38594*	**9,5**
Deutschland	*32962*	**7,3**
Dominikanische Republik	*2062*	**3,5**
Ecuador	*1425*	**2,2**

El Salvador	*1725*	**3,4**
Estland	*3279*	**5,6**
Finnland	*31980*	**9,7**
Frankreich	*29807*	**6,3**
Georgien	*499*	**2,4**
Ghana	*413*	**3,9**
Griechenland	*13192*	**4,2**
Guatemala	*1558*	**2,5**
Haiti	*367*	**2,2**
Honduras	*711*	**2,7**
Indien	*459*	**2,7**
Indonesien	*994*	**1,9**
Irland	*29087*	**6,9**
Island	*31602*	**9,4**
Israel	*17067*	**7,3**
Italien	*20860*	**5,2**
Jamaika	*1785*	**4,0**
Japan	*44695*	**7,1**
Jordanien	*1616*	**4,5**
Kamerun	*675*	**2,2**
Kanada	*22925*	**9,0**
Kasachstan	*1512*	**2,3**
Kenia	*328*	**1,9**
Kolumbien	*2290*	**3,6**
Korea, Republik	*13212*	**4,5**
Kroatien	*5146*	**3,8**
Lettland	*2314*	**3,7**
Litauen	*1920*	**4,8**
Luxemburg	*56110*	**9,0**
Madagaskar	*246*	**1,7**
Malawi	*169*	**2,9**
Malaysia	*4797*	**4,9**
Marokko	*1370*	**3,7**
Mauritius	*4429*	**4,5**
Mexiko	*3787*	**3,6**
Moldau, Republik	*636*	**2,1**

Namibia	*2408*	**5,7**
Neuseeland	*18200*	**9,5**
Nicaragua	*466*	**2,5**
Niederlande	*31199*	**9,0**
Nigeria	*254*	**1,6**
Norwegen	*38142*	**8,5**
Österreich	*32919*	**7,8**
Pakistan	*516*	**2,6**
Panama	*3279*	**3,0**
Paraguay	*1700*	**1,7**
Peru	*2368*	**4,0**
Philippinen	*1167*	**2,6**
Polen	*4231*	**4,0**
Portugal	*12664*	**6,3**
Rumänien	*1479*	**2,6**
Russische Föderation	*2455*	**2,7**
Sambia	*392*	**2,6**
Schweden	*31329*	**9,3**
Schweiz	*46821*	**8,5**
Senegal	*609*	**3,1**
Simbabwe	*621*	**2,7**
Singapur	*28230*	**9,3**
Slowakei	*4162*	**3,7**
Slowenien	*11659*	**6,0**
Spanien	*17633*	**7,1**
Sri Lanka	*860*	**3,7**
Südafrika	*3985*	**4,8**
Taiwan	*734*	**5,6**
Tansania, Vereinigte Republik	*190*	**2,7**
Thailand	*2805*	**3,2**
Trinidad und Tobago	*5123*	**4,9**
Tschechische Republik	*5312*	**3,7**
Tunesien	*2470*	**4,8**
Türkei	*3076*	**3,2**
Uganda	*348*	**2,1**
Ukraine	*896*	**2,4**

Ungarn	*5459*	**4,9**
Uruguay	*6115*	**5,1**
Usbekistan	*485*	**2,9**
Venezuela	*3300*	**2,5**
Vereinigte Staaten	*31730*	**7,7**
Vereinigtes Königreich	*21568*	**8,7**
Vietnam	*356*	**2,4**
Weißrussland	*2760*	**4,8**

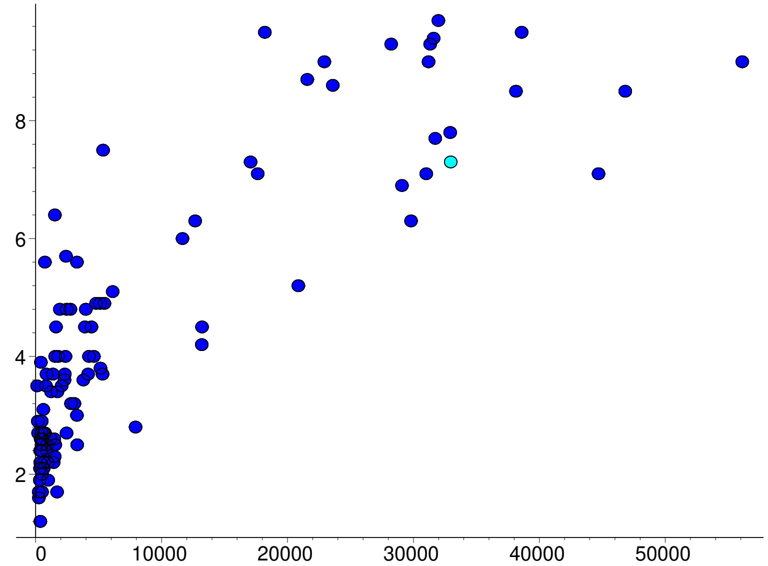

Abb. 18: Zusammenhang zwischen Korruption und Bruttoinlandsprodukt je Einwohner

Abbildung 18 zeigt die 101 Datenpunkte zu dieser Liste graphisch, wobei der Punkt für Deutschland etwas heller eingezeichnet ist.

Der erste Augenschein zeigt, daß korruptionsärmere Länder oftmals reicher sind: Das weitgehend korruptionsfreie Dänemark hat ein Bruttosozialprodukt von 38 594 \$ pro Einwohner, das deutlich korruptere Deutschland nur 32 962 \$ und ein stark korruptes Land wie Bangladesch nur 373 \$. Allerdings gibt es auch Ausnahmen, denn Chile hat,

obwohl geringfügig weniger korrupt als Deutschland, nur einBruttosozialprodukt von 5 354 \$ pro Einwohner. Es gibt also sicherlich keinen deterministischen Zusammenhang zwischen Korruption und Wohlstand, aber doch eine Tendenz.

Falls wir nun versuchen, beispielsweise einen linearen Zusammenhang der Form

$$CPI = a + b \cdot BIP$$

zu finden, so haben wir 101 Gleichungen für die beiden unbekannten Koeffizienten a und b, und ein kurzer Blick auf Abbildung 18 zeigt, daß dieses lineare Gleichungssystem keine Lösung haben kann.

Wir suchen also keine Lösung, sondern zwei Zahlen a und b derart, daß die 101 Gleichungen „möglichst gut" gelten. Was das bedeuten soll läßt sich mathematisch auf verschiedene, nicht äquivalente Weisen definieren; da wir uns im Augenblick mit Skalarprodukten beschäftigen, bietet sich an, die 101 Bruttoinlandsprodukte pro Einwohner und die 101 Korruptionsindizes zu zwei Vektoren $\vec{x}, \vec{y} \in \mathbb{R}^{101}$ zusammenzufassen, und nach Zahlen a, b zu suchen, so daß die Länge des Differenzvektors $\vec{y} - a\vec{x} - b$ möglichst klein wird. Ausgeschrieben bedeutet dies, wenn wir die Komponenten von $\vec{x}$ mit x_i und die von $\vec{y}$ mit y_i bezeichnen, daß die Summe

$$\sum_{i=1}^{101} (y_i - ax_i - b)^2$$

der Abweichungsquadrate möglichst klein sein soll – von daher der Name „Methode der kleinsten Quadrate" für diesen Ansatz, mit dessen Hilfe sein Schöpfer GAUSS sowohl die Position des Planetoiden Ceres vorhersagte als auch die Vermessung und Kartierung des Königreichs Hannover durchführte.

Derselbe Ansatz läßt sich natürlich auf jedes lineare Gleichungssystem über den reellen oder komplexen Zahlen anwenden: Wir haben ein möglicherweise unlösbares lineares Gleichungssystem $A\vec{x} = \vec{b}$ und wollen einen Vektor $\vec{x}$ so bestimmen, daß der Vektor $A\vec{x} - \vec{b}$ minimale Länge hat.

Falls das lineare Gleichungssystem lösbar ist, gibt es damit kein Problem: Wir bestimmen irgendeine Lösung $\vec{x}$ und haben damit einen Vektor gefunden, für den $A\vec{x}-\vec{b}$ die Länge null hat – kürzer geht es nicht.

Im allgemeinen ist aber für den gesuchten Vektor $\vec{x}$ das Produkt $A\vec{x}$ von $\vec{b}$ verschieden; es sei etwa gleich $\vec{c}$. Dann ist $\vec{c}$ ein Vektor, der sich in der Form $A\vec{x}$ darstellen läßt, und unter allen solchen Vektoren ist es derjenige, für den die Länge des Differenzvektors zu $\vec{b}$ minimal ist. Dies erinnert an die orthogonalen Projektionen aus dem vorigen Abschnitt, und in der Tat läßt sich das Problem damit lösen:

Nehmen wir an, wir haben n Gleichungen in m Unbekannten mit Koeffizienten aus $k=\mathbb{R}$ oder $k=\mathbb{C}$. Dann definiert die Matrix $A\in k^{n\times m}$ des Gleichungssystems eine lineare Abbildung

$$\varphi\colon k^m\to k^n;\quad \vec{v}\mapsto A\vec{v};$$

deren Bildraum sei U. Falls die rechte Seite $\vec{b}$ in U liegt, ist das Gleichungssystem lösbar; andernfalls suchen wir einen Vektor $\vec{x}\in k^m$, für den die Länge des Vektors $A\vec{x}-\vec{b}$ minimal wird. Da die Vektoren, die sich in der Form $A\vec{x}$ darstellen lassen, genau die Vektoren aus U sind, ist somit $A\vec{x}=\pi_U(\vec{b})$ die orthogonale Projektion von $\vec{b}$ nach U. Diese könnten wir *im Prinzip* bestimmen, indem wir die QR-Zerlegung von A berechnen, denn dann sind die ersten Spalten von Q eine Basis von U, die durch die weiteren Spalten zu einer Basis von ganz k^n ergänzt wird; danach haben wir ein lösbares lineares Gleichungssystem.

Wir wollen uns überlegen, wie wir $\vec{x}$ auch ohne die rechnerisch aufwendige QR-Zerlegung bestimmen können.

Für den gesuchten Vektor $\vec{x}$ (oder für die gesuchten Vektoren $\vec{x}$) ist $A\vec{x}=\varphi_U(\vec{b})$. Da $A\vec{x}$ bereits in U liegt, ist $\pi_U(A\vec{x})=A\vec{x}$, also ist die Gleichung $A\vec{x}=\pi_U(\vec{b})$ äquivalent zu

$$\pi_U(A\vec{x})=\pi(U)(\vec{b})\quad\text{oder}\quad A\vec{x}-\vec{b}\in\operatorname{Kern}\pi_U=U^\perp\,.$$

Das orthogonale Komplement $U^\perp$ von U besteht aus allen Vektoren $\vec{y}\in k^n$, die senkrecht stehen auf U, für die also gilt

$$(A\vec{x})\cdot\vec{y}=0\quad\text{für alle }\vec{x}\in k^m\,.$$

Wie wir im vorletzten Abschnitt gesehen haben, ist

$$(A\vec{x})\cdot\vec{y}=\vec{x}\cdot A^*\vec{y}\quad\text{für alle }\vec{x}\in k^m,\vec{y}\in k^n\,,$$

$\vec{y}$ liegt also genau dann in $U^\perp$, wenn $A^*\vec{y}$ senkrecht steht auf allen Vektoren $\vec{x}\in k^m$. Ein solcher Vektor aus k^m ist insbesondere $A^*\vec{y}$ selbst; wegen der positiven Definitheit des (HERMITEschen) Skalarprodukts ist also $A^*\vec{y}=\vec{0}$ und somit

$$U^\perp=\left\{\vec{y}\in k^n\;\middle|\;A^*\vec{y}=\vec{0}\right\}.$$

$A\vec{x}-\vec{b}$ liegt also genau dann im Kern von π_U, wenn $A^*(A\vec{x}-\vec{b})=\vec{0}$ ist oder, anders ausgedrückt, wenn $\vec{x}$ eine Lösung des linearen Gleichungssystems

$$(A^*A)\,\vec{x}=A^*\vec{b}$$

ist. Da die adjungierte Matrix A^* einfach die transponierte Matrix zur komplex konjugierten Matrix zu A ist, wobei die komplexe Konjugation über $\mathbb{R}$ natürlich entfällt, läßt sich dieses Gleichungssystem schnell aufstellen und dann nach GAUSS lösen.

Betrachten wir dies konkret im eingangs diskutierten Fall eines linearen Zusammenhangs $y=ax+b$ zu N Wertepaaren $(x_i,y_i)\in\mathbb{R}^2$, wobei N sinnvollerweise größer als zwei sein sollte. Wir haben dann N Gleichungen

$$y_i=ax_i+b\quad\text{oder}\quad x_ia+b=y_i\,,$$

wobei hier im Gegensatz zu unserer sonstigen Gewohnheit die Parameter a und b unbekannt sind, während die x_i und die y_i bekannt sind. Wir haben also ein lineares Geichungssystem von N Gleichungen in den beiden Variablen a und b.

Fassen wir die Werte x_i zusammen zu einem Vektor $\vec{x}\in\mathbb{R}^N$ und die y_i zu einem Vektor $\vec{y}\in\mathbb{R}^n$, so läßt sich dieses Gleichungssystem kurz schreiben als

$$\vec{x}\cdot a+\vec{1}\cdot b=\vec{y}\,,$$

wobei $\vec{1}\in\mathbb{R}^N$ jenen Vektor bezeichnen soll, dessen sämtliche Komponenten eins sind.

Die Matrix des Gleichungssystems ist somit die $N \times 2$-Matrix A mit Spalten $\vec{x}$ und $\vec{1}$. Da wir mit reellen Zahlen rechnen, ist A^* einfach die transponierte Matrix dazu, also die $2 \times N$-Matrix, in deren erster Zeile die x_i stehen, während in der zweiten lauter Einser stehen. Somit ist

$$ {}^tAA = \begin{pmatrix} \vec{x}\cdot\vec{x} & \vec{x}\cdot\vec{1} \\ \vec{x}\cdot\vec{1} & \vec{1}\cdot\vec{1} \end{pmatrix} \quad \text{und} \quad {}^tA\vec{b} = \begin{pmatrix} \vec{x}\cdot\vec{y} \\ \vec{1}\cdot\vec{y} \end{pmatrix}, $$

das Gleichungssystem wird also zu

$$ (\vec{x}\cdot\vec{x})a + (\vec{x}\cdot\vec{1})b = \vec{x}\cdot\vec{y} \quad \text{und} \quad (\vec{x}\cdot\vec{1}) + Nb = \vec{y}\cdot\vec{1}\,. $$

Seine Matrix ist genau dann singulär, wenn die beiden Spalten proportional zueinander sind, wenn also $N(\vec{x}\cdot\vec{x}) = (\vec{x}\cdot\vec{1})^2$ ist. Nach der CAUCHY-SCHWARZschen Ungleichung ist

$$ \left|\vec{1}\cdot\vec{x}\right| \le \left|\vec{1}\right| \cdot |\vec{x}| = \sqrt{N}\,|\vec{x}|\,, \quad \text{also} \quad \left|\vec{1}\cdot\vec{x}\right|^2 \le N(\vec{x}\cdot\vec{x}) $$

mit Gleichheit nur dann, wenn die Vektoren $\vec{x}$ und $\vec{1}$ linear abhängig sind, wenn also alle x_i gleich sind. In diesem Fall ist die erste Gleichung ein Vielfaches der zweiten, es gibt also unendlich viele Lösungen.

Andernfalls ist die Matrix invertierbar, die Lösung also eindeutig.

Führen wir die (in der Ausgleichsrechnung ziemlich verbreiteten) Abkürzungen

$$ [x^r] = \sum_{i=1}^{N} x_i^r\,, \quad [y^r] = \sum_{i=1}^{N} x_i^r \quad \text{und} \quad [x^r y^s] = \sum_{i=1}^{N} x_i^r y_i^s $$

ein, so erhält das Gleichungssystem die übersichtlichere Gestalt

$$ [x^2]a + [x]b = [xy] \quad \text{und} \quad [x]a + Nn = [y]\,. $$

Subtraktion von $[x]/[x^2]$ mal der ersten Gleichung von der zweiten führt auf

$$ \left(N - \frac{[x]^2}{[x^2]}\right) b = [y] - \frac{[x]}{[x^2]}[xy] $$

oder $(N[x^2] - [x]^2)b = [y][x^2] - [x][xy]$, d.h.

$$ b = \frac{[y][x^2] - [x][xy]}{N[x^2] - [x]^2}\,. $$

(Man beachte, daß im Falle der eindeutigen Lösbarkeit sowohl $[x^2] > 0$ als auch $N[x^2] - [x]^2 > 0$ ist.)

Einsetzen von b in die erste Gleichung ergibt dann auch

$$ a = \frac{[xy] - [x]b}{[x^2]}\,. $$

Im Falle des Zusammenhangs zwischen Korruptionsindex CPI und Bruttoinlandsprodukt pro Einwohner BIP erhalten wir nach diesen Formeln die Ausgleichsgerade

$$ CPI = 3{,}145 + 0{,}000153\,BIP\,, $$

die Steigung ist also erwartungsgemäß positiv. Der relativ große konstante Term zeigt, daß *im Mittel* Korruption selbst bei sehr armen Ländern deutlich über dem unteren Ende der Skala liegt.

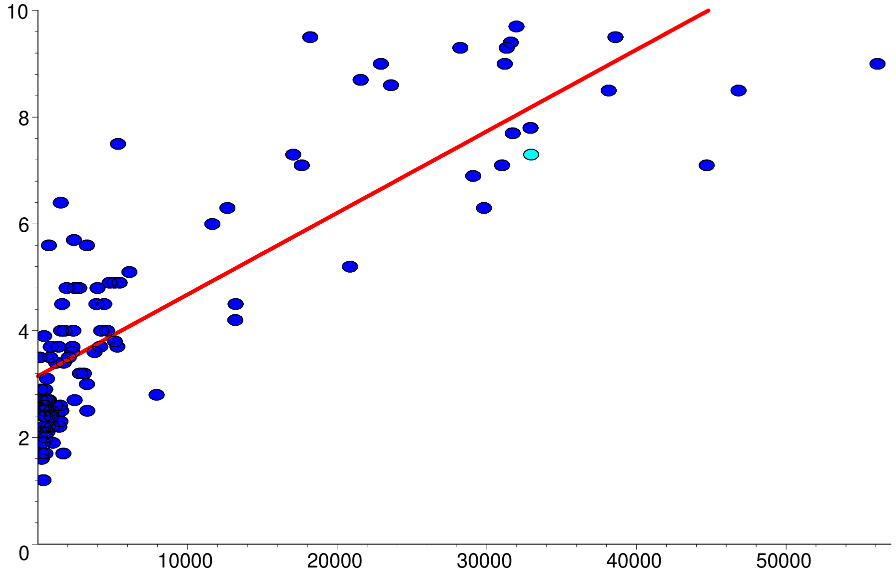

Abb. 19: Ausgleichsgerade zu Abbildung 18