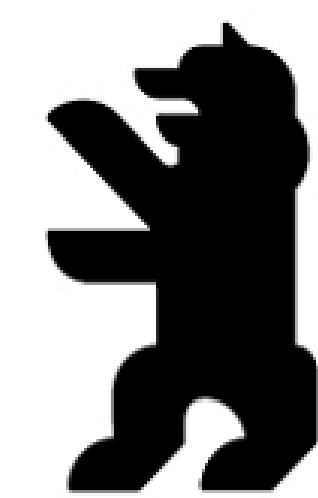


The Geometry of Moore-Penrose Generalized Matrix Inverses

Martin Erik Horn

BSEL/HWR Berlin, FB 1 – Department of Business and Economics, FE Quantitative Methods



Hochschule für
Wirtschaft
und Recht Berlin

Berlin School of Economics and Law

Email: e_hornm@doz.hwr-berlin.de

Economathematical Starting Point

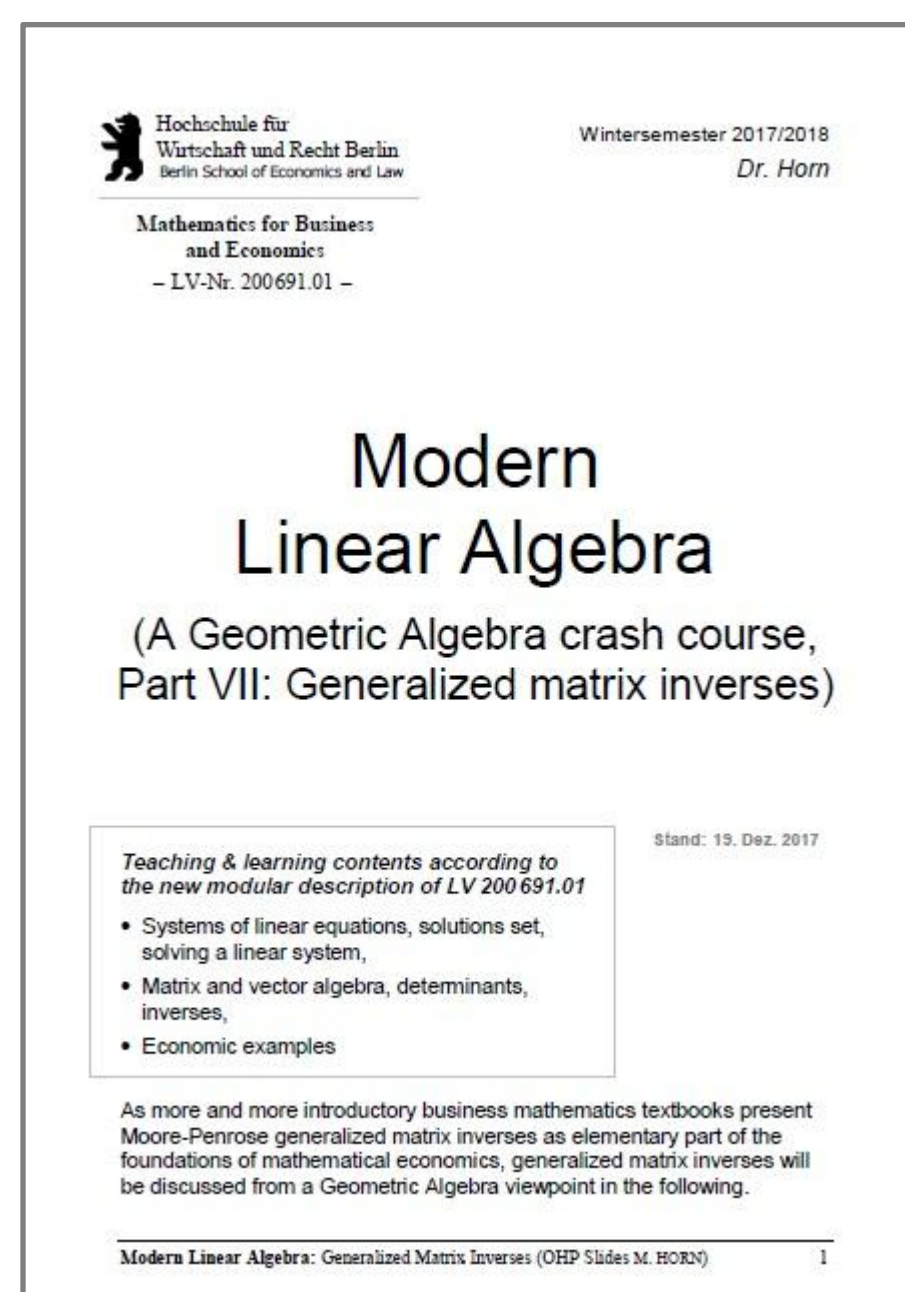
More and more introductory business mathematics textbooks present Moore-Penrose generalized matrix inverses as elementary part of the foundations of mathematical economics. Generalized matrix inverses are regularly discussed in introductory courses e.g. at FH Schmalkalden & TU Dortmund.

Didactical Problem

Most textbooks introduce generalized matrix inverses by purely algebraic reasoning and the discussion of Moore-Penrose generalized matrix inverses is based on the four Moore-Penrose conditions:

$$\begin{aligned} \mathbf{A} \mathbf{A}^+ \mathbf{A} &= \mathbf{A} & (\mathbf{A}^+ \mathbf{A})^T &= \mathbf{A}^+ \mathbf{A} \\ \mathbf{A}^+ \mathbf{A} \mathbf{A}^+ &= \mathbf{A}^+ & (\mathbf{A} \mathbf{A}^+)^T &= \mathbf{A} \mathbf{A}^+ \end{aligned}$$

But to give a complete picture of these mathematical structures it is helpful to introduce and to describe Moore-Penrose generalized matrix inverses also by using geometric representations based on the ideas of Grassmann's theory of extensions.



After having discussed the basics of Geometric Algebra and the Geometric Algebra solution scheme of systems of linear equations with the students at previous lessons, a two hour lecture (2 x 45 min.) was required to introduce Pauli algebra generalized matrix inverses and Moore-Penrose generalized matrix inverses.

Overview of BSEL Geometric Algebra Crash Course:

- Part 1: Basic Foundations of Geometric Algebra
 - Part 2: Solving Systems of Linear Equations with two or three Unknown Variables
 - Part 3: Direct Product & Solving higher-dimensional Systems of Linear Equations
 - Part 4: Transformation of Coordinates & Gaussian Method of Solving Systems of Linear Equations
 - Part 5: Eigenvalues and Eigenvectors
 - Part 6: Solving Systems of Linear Equations with Sandwich Products
 - Part 7: Generalized Matrix Inverses
 - Addendum: Solving Systems of Linear Equations with the Geometric Algebra Algorithms Optimizer (GAALOP)
- winter semester 2014/2015**
 → DPG Wuppertal 2015
- winter semester 2015/2016**
 → DPG Hannover 2016
- winter semester 2016/2017**
 → DPG Dresden 2017
- winter semester 2017/2018**
 → DPG Würzburg & Berlin 2018

Second Starting Point from the Perspective of Physics: Pauli Algebra and Generalized Pauli Algebra (Geometric Algebra)

Pauli Matrices represent base vectors of three-dimensional space.
Generalized Pauli Matrices represent base vectors of higher-dimensional spaces.

(And Dirac Matrices represent base vectors of four- or five-dimensional spacetimes.)

→ More about the foundation of this perspective will be discussed at the short talk SOE 9.2 "Pauli Algebras in Economics: Economathematics from Geometry to Didactics and back" at Tuesday, March 13, 2018, 10:45 – 11:00 h in room MA 001.

Die Wissenschaft
der
extensiven Grösse
oder
die Ausdehnungslehre,
eine neue mathematische Disciplin
dargestellt und durch Anwendungen erläutert
von
Hermann Grassmann
Lehrer an der Friedrich-Wilhelms-Schule in Stettin
Erster Theil,
die lineale Ausdehnungslehre enthaltend.
Leipzig, 1844.
Verlag von Otto Wigand

§ 45. Dass nun die äussere Multiplikation, da sie den Begriff des Verschiedenartigen wesentlich voraussetzt, auf die Zahlenlehre keine so unmittelbare Anwendung findet, wie auf die Geometrie und Mechanik, darf uns freilich nicht wundern, indem die Zahlen ihrem Inhalte nach als gleichartige erscheinen. Aber desto interessanter ist es, zu bemerken, wie in der Algebra, sobald an der Zahl noch die Art ihrer Verknüpfung mit andern Grössen festgehalten, und in dieser Hinsicht die eine als von der andern formell verschiedenartig aufgefasst wird, auch die Anwendbarkeit der äusseren Multiplikation mit einer so schlagenden Entscheidung heranstritt, dass ich wohl behaupten darf, es werde durch diese Anwendung auch die Algebra eine wesentlich veränderte Gestalt gewinnen. Um hieron eine Idee zu geben, will ich n Gleichungen ersten Grades mit n Unbekannten setzen, von der Form

$$\begin{aligned} a_1 x_1 + a_2 x_2 + \dots + a_n x_n &= b_0 \\ b_1 x_1 + b_2 x_2 + \dots + b_n x_n &= b_1 \\ \dots & \dots \\ s_1 x_1 + s_2 x_2 + \dots + s_n x_n &= s_0 \end{aligned}$$

wo x_1, \dots, x_n die Unbekannten seien. Hier können wir die Zahlenkoeffizienten, welche verschiedenen Gleichungen angehören, sofern wir diese Verschiedenheit an ihrem Begriff noch festhalten, als verschiedenartig ansehen, und zwar alle als an sich verschiedenartig, d. h. als unabhängig in dem Sinne unserer Wissenschaft, die einer und derselben Gleichung als unter sich in derselben Beziehung gleichartig. Addiren wir nun in diesem Sinne alle n Gleichungen und bezeichnen die Summe des Verschiedenartigen in dem Sinne unserer Wissenschaft mit dem Verknüpfungszeichen \wedge ,

indem die gleichen Stellen in den so gebildeten Summenausdrücken immer dem Gleichartigen zukommen sollen, so erhalten wir

$$(a_1 + b_1 + \dots + s_1)x_1 + (a_2 + b_2 + \dots + s_2)x_2 + \dots + (a_n + b_n + \dots + s_n)x_n = (b_0 + b_1 + \dots + s_0)$$

oder bezeichnen wir $(a_1 + b_1 + \dots + s_1)$ mit p_1 und entsprechend die übrigen Summen, so haben wir

$$p_1 x_1 + p_2 x_2 + \dots + p_n x_n = p_0$$

Aus dieser Gleichung, welche die Stelle jener n Gleichungen vertritt, lässt sich nun auf der Stelle jeder der Unbekannten, z. B. x_1 finden, wenn wir die beiden Seiten mit dem äusseren Produkte aus den Koeffizienten der übrigen Unbekannten äusserlich multiplizieren, also hier mit $p_2 \cdot p_3 \cdot \dots \cdot p_n$. Da nämlich, wenn man die Glieder der linken Seite einzeln multipliziert, nach dem Begriff des äusseren Produktes (§ 31) alle Produkte wegfällt, welche zwei gleiche Faktoren enthalten, so erhält man

$$p_1 \cdot p_2 \cdot p_3 \cdot \dots \cdot p_n x_1 = p_0 \cdot p_2 \cdot p_3 \cdot \dots \cdot p_n$$

Also da beide Produkte, als denselben System n-ter Stufe angehörig einander gleichartig sind, so hat man

$$x_1 = \frac{p_0 \cdot p_2 \cdot p_3 \cdot \dots \cdot p_n}{p_1 \cdot p_2 \cdot p_3 \cdot \dots \cdot p_n}$$

Also jede Unbekannte ist einem Bruche gleich, dessen Nenner das äussere Produkt der Koeffizienten p_2, \dots, p_n ist, und dessen Zähler man erhält, wenn man in diesem Produkt statt des Koeffizienten jener Unbekannten die rechte Seite, nämlich p_0 , als Faktor setzt. Alle Unbekannten haben also denselben Nenner, und werden unbestimmt oder unendlich, wenn dieser Nenner null wird, d. h. ist.

$$p_1 \cdot p_2 \cdot p_3 \cdot \dots \cdot p_n = 0$$

ist.

With his theory of extensions Herman Günther Grassmann (1809 – 1877) already invented generalized Pauli Algebra and generalized Dirac Algebra. The solution of a system of linear equations can be found by applying his solution equations of the first edition of his **Ausdehnungslehre** of 1844.

Written in modern form, the solution of consistent systems of linear equations with two variables x, y can be found with the following solution equations:

$$\mathbf{a}x + \mathbf{b}y = \mathbf{r}$$

$$x = (\mathbf{a} \wedge \mathbf{b})^{-1} (\mathbf{r} \wedge \mathbf{b}) \quad \text{or} \quad x = (\mathbf{r} \wedge \mathbf{b}) (\mathbf{a} \wedge \mathbf{b})^{-1}$$

$$y = (\mathbf{a} \wedge \mathbf{b})^{-1} (\mathbf{a} \wedge \mathbf{r}) \quad \text{or} \quad y = (\mathbf{a} \wedge \mathbf{r}) (\mathbf{a} \wedge \mathbf{b})^{-1}$$

Geometric Interpretation of the Solution Equations

Outer products of two vectors can be interpreted as oriented area elements. And as a change of the coordinate system does not change the geometric situation, all ratios of the areas of the oriented parallelograms $\mathbf{a} \wedge \mathbf{b}$, $\mathbf{r} \wedge \mathbf{b}$, and $\mathbf{a} \wedge \mathbf{r}$ do not depend on the coordinate system.

coefficient vectors:
 $\mathbf{a} = a_1 \sigma_x + a_2 \sigma_y$
 $\mathbf{b} = b_1 \sigma_x + b_2 \sigma_y$
 resulting vector:
 $\mathbf{r} = r_1 \sigma_x + r_2 \sigma_y$

coefficient vectors:
 $\mathbf{a} = a_1 \sigma_x + a_2 \sigma_y + a_3 \sigma_z$
 $\mathbf{b} = b_1 \sigma_x + b_2 \sigma_y + b_3 \sigma_z$
 resulting vector:
 $\mathbf{r} = r_1 \sigma_x + r_2 \sigma_y + r_3 \sigma_z$

But a change of the coordinate system will change the algebraic description of the vectors, which now have three components instead of only two. Therefore the matrix inverse will have a different algebraic description, too. This generalized Pauli Algebra matrix inverse can be used to solve **consistent** systems of linear equations, even if these systems have more equations than variables.

Conventional Matrix Inverses and Generalized Matrix Inverses

As Grassmann's solution equations can always be written as a matrix multiplication, the lead matrix of this matrix multiplication will be the matrix inverse:

$$\begin{aligned} x &= (\mathbf{a} \wedge \mathbf{b})^{-1} (\mathbf{r} \wedge \mathbf{b}) = (\mathbf{a} \wedge \mathbf{b})^{-1} ((\sigma_x \wedge \mathbf{b}) r_1 + (\sigma_y \wedge \mathbf{b}) r_2) \\ y &= (\mathbf{a} \wedge \mathbf{b})^{-1} (\mathbf{a} \wedge \mathbf{r}) = (\mathbf{a} \wedge \mathbf{b})^{-1} ((\mathbf{a} \wedge \sigma_x) r_1 + (\mathbf{a} \wedge \sigma_y) r_2) \end{aligned}$$

$$\begin{aligned} x &= (\mathbf{a} \wedge \mathbf{b})^{-1} (\mathbf{r} \wedge \mathbf{b}) = (\mathbf{a} \wedge \mathbf{b})^{-1} ((\sigma_x \wedge \mathbf{b}) r_1 + (\sigma_y \wedge \mathbf{b}) r_2 + (\sigma_z \wedge \mathbf{b}) r_3) \\ y &= (\mathbf{a} \wedge \mathbf{b})^{-1} (\mathbf{a} \wedge \mathbf{r}) = (\mathbf{a} \wedge \mathbf{b})^{-1} ((\mathbf{a} \wedge \sigma_x) r_1 + (\mathbf{a} \wedge \sigma_y) r_2 + (\mathbf{a} \wedge \sigma_z) r_3) \end{aligned}$$

Inverse of a 2 x 2 square matrix: $\mathbf{A}^{-1} = \frac{1}{\mathbf{a} \wedge \mathbf{b}} \begin{bmatrix} \sigma_x \wedge \mathbf{b} & \sigma_y \wedge \mathbf{b} \\ \mathbf{a} \wedge \sigma_x & \mathbf{a} \wedge \sigma_y \end{bmatrix}$

Inverse of a 3 x 2 matrix: $\mathbf{A}^{-1} = \frac{1}{\mathbf{a} \wedge \mathbf{b}} \begin{bmatrix} \sigma_x \wedge \mathbf{b} & \sigma_y \wedge \mathbf{b} & \sigma_z \wedge \mathbf{b} \\ \mathbf{a} \wedge \sigma_x & \mathbf{a} \wedge \sigma_y & \mathbf{a} \wedge \sigma_z \end{bmatrix}$
 (if the bivector $\mathbf{a} \wedge \mathbf{b}$ is pre-multiplied from the left)

If $(\mathbf{a} \wedge \mathbf{b}) \neq 0$, every element of \mathbf{A}^{-1} will be a scalar.

Or alternatively: $\mathbf{A}^{-1} = \begin{bmatrix} \sigma_x \wedge \mathbf{b} & \sigma_y \wedge \mathbf{b} & \sigma_z \wedge \mathbf{b} \\ \mathbf{a} \wedge \sigma_x & \mathbf{a} \wedge \sigma_y & \mathbf{a} \wedge \sigma_z \end{bmatrix} \frac{1}{\mathbf{a} \wedge \mathbf{b}}$
 (if the bivector $\mathbf{a} \wedge \mathbf{b}$ is post-multiplied from the right)

Conclusion:

Moore-Penrose generalized matrix inverses \mathbf{A}^+ consist of the scalar terms of Pauli algebra generalized matrix inverses \mathbf{A}^{-1} , which usually possess higher-dimensional terms, too.

As all elements of \mathbf{A}^{-1} are products of two different bivectors, every element will be a linear combination of a scalar and a bivector.

$$\mathbf{A}^+ = \begin{bmatrix} \langle (\mathbf{a} \wedge \mathbf{b})^{-1} (\sigma_x \wedge \mathbf{b}) \rangle_0 & \langle (\mathbf{a} \wedge \mathbf{b})^{-1} (\sigma_y \wedge \mathbf{b}) \rangle_0 & \langle (\mathbf{a} \wedge \mathbf{b})^{-1} (\sigma_z \wedge \mathbf{b}) \rangle_0 \\ \langle (\mathbf{a} \wedge \mathbf{b})^{-1} (\mathbf{a} \wedge \sigma_x) \rangle_0 & \langle (\mathbf{a} \wedge \mathbf{b})^{-1} (\mathbf{a} \wedge \sigma_y) \rangle_0 & \langle (\mathbf{a} \wedge \mathbf{b})^{-1} (\mathbf{a} \wedge \sigma_z) \rangle_0 \end{bmatrix}$$

→ Pauli Algebra generalized matrix inverses are left-sided matrix inverses:

$$\mathbf{A}^{-1} \mathbf{A} = \mathbf{I} \neq \mathbf{A} \mathbf{A}^{-1}$$

$$\begin{aligned} \mathbf{A} \mathbf{A}^{-1} \mathbf{A} &= \mathbf{A} \\ \mathbf{A}^{-1} \mathbf{A} \mathbf{A}^{-1} &= \mathbf{A}^{-1} \end{aligned}$$

$$\begin{aligned} (\mathbf{A}^{-1} \mathbf{A})^T &= \mathbf{A}^{-1} \mathbf{A} \\ (\mathbf{A} \mathbf{A}^{-1})^T &\neq \mathbf{A} \mathbf{A}^{-1} \end{aligned}$$

The fourth Moore-Penrose condition is only met by the scalar terms of \mathbf{A}^{-1} .

$$(\mathbf{A} \langle \mathbf{A}^{-1} \rangle_0)^T = \mathbf{A} \langle \mathbf{A}^{-1} \rangle_0$$

