Harjoitusten 1 ratkaisuehdotukset

1. Häivytetään merkintöjen yksinkertaistamiseksi muuttujien riippuvuus $\boldsymbol{\beta}$:sta.

   a) Todistus on suoraviivainen:

$$
\begin{aligned}
\frac{\partial(\mathbf{a}'\mathbf{A}\mathbf{c})}{\partial\boldsymbol{\beta}'} &= \frac{\partial}{\partial\boldsymbol{\beta}'}(\mathbf{a}'\begin{bmatrix} \sum_{j=1}^{p} a_{1j}c_j \\ \vdots \\ \sum_{j=1}^{p} a_{nj}c_j \end{bmatrix}) \\
&= \frac{\partial}{\partial\boldsymbol{\beta}'}(\textstyle\sum_{i=1}^{n}\sum_{j=1}^{p} a_i a_{ij} c_j) \\
&= \textstyle\sum_{i=1}^{n}\sum_{j=1}^{p}(\displaystyle\frac{\partial a_i}{\partial\boldsymbol{\beta}'} a_{ij}c_j + a_i a_{ij}\frac{\partial c_j}{\partial\boldsymbol{\beta}'}) \\
&= \textstyle\sum_{j=1}^{p} c_j \sum_{i=1}^{n} a_{ij}\displaystyle\frac{\partial a_i}{\partial\boldsymbol{\beta}'} + \textstyle\sum_{i=1}^{n} a_i \sum_{j=1}^{p} a_{ij}\displaystyle\frac{\partial c_j}{\partial\boldsymbol{\beta}'}) \\
&= \textstyle\sum_{j=1}^{p} c_j \begin{bmatrix} \sum_{i=1}^{n} a_{ij}\frac{\partial a_i}{\partial\beta_1} & \cdots & \sum_{i=1}^{n} a_{ij}\frac{\partial a_i}{\partial\beta_m} \end{bmatrix} \\
&\quad + \textstyle\sum_{i=1}^{n} a_i \begin{bmatrix} \sum_{j=1}^{p} a_{ij}\frac{\partial c_j}{\partial\beta_1} & \cdots & \sum_{j=1}^{p} a_{ij}\frac{\partial c_j}{\partial\beta_m} \end{bmatrix} \\
&\quad \text{[huomaa että ylemmällä rivillä "juoksee" "}i\text{" ja alemmalla "}j\text{"]} \\
&= \mathbf{c}'\mathbf{A}'\begin{bmatrix} \frac{\partial a_1}{\partial\beta_1} & \cdots & \frac{\partial a_1}{\partial\beta_m} \\ \vdots & & \vdots \\ \frac{\partial a_n}{\partial\beta_1} & \cdots & \frac{\partial a_n}{\partial\beta_m} \end{bmatrix} + \mathbf{a}'\mathbf{A}\begin{bmatrix} \frac{\partial c_1}{\partial\beta_1} & \cdots & \frac{\partial c_1}{\partial\beta_m} \\ \vdots & & \vdots \\ \frac{\partial c_p}{\partial\beta_1} & \cdots & \frac{\partial c_p}{\partial\beta_m} \end{bmatrix} \\
&= \mathbf{c}'\mathbf{A}'\frac{\partial\mathbf{a}}{\partial\boldsymbol{\beta}'} + \mathbf{a}'\mathbf{A}\frac{\partial\mathbf{c}}{\partial\boldsymbol{\beta}'}.
\end{aligned}
$$

b) Väitteet eivät seuraa suoraan a)-kohdasta. Siinä derivoidaan skalaari, ja tulos on $1 \times m$-vektori. Tämän kohdan tehtävissä derivoidaan $m \times 1$ -vektori ja tulos on $m \times n$ -matriisi. Myöskään merkinnät eivät täsmäisi, jos esimerkiksi $n \times 1$-vektori $\mathbf{a}(\boldsymbol{\beta})$ asetettaisiin skalaariksi. Tällöin merkintöjen mukaan $n \times p$ -matriisi $\mathbf{A}$ typistyisi $1 \times p$-vektoriksi.[1]

Väite on kuitenkin johdettavissa a)-kohdan perusteella. Asetetaan siinä $\mathbf{a}(\boldsymbol{\beta}) = [0 \ldots 0 \quad 1 \quad 0 \ldots 0]' \equiv \mathbf{e}_t$ $(m \times 1)$, jossa "1" on t. termi $(t = 1, \ldots m)$ ja $\mathbf{c}(\boldsymbol{\beta}) = \boldsymbol{\beta}$. Tällöin $\partial \mathbf{a}/\partial \boldsymbol{\beta}' = \mathbf{0}_{m\times n}$ (nollista koostuva $m \times n$ -matriisi) ja $\partial \mathbf{c}/\partial \boldsymbol{\beta}' = \mathbf{I}_n$. Derivoitava muuttuja on näillä asetuksilla $\mathbf{e}_t^{'}\mathbf{A}\boldsymbol{\beta} = \mathbf{A}_t\boldsymbol{\beta}$, jossa $\mathbf{A}_t$ on $\mathbf{A}$:n t's rivi $(1 \times n)$:

$$\begin{aligned}
\frac{\partial(\mathbf{A}_t\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'} &= \frac{\partial(\mathbf{e}_t^{'}\mathbf{A}\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'} \\
&= \left[\frac{\partial \mathbf{a}(\boldsymbol{\beta})'\mathbf{A}\mathbf{c}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'}\right]\Bigg|_{\mathbf{a}(\boldsymbol{\beta})=\mathbf{e}_t,\ \mathbf{c}(\boldsymbol{\beta})=\boldsymbol{\beta}} \\
&\overset{(2)}{=} \left[\mathbf{c}(\boldsymbol{\beta})'\mathbf{A}'\frac{\partial\mathbf{a}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'} + \mathbf{a}(\boldsymbol{\beta})'\mathbf{A}\frac{\partial\mathbf{c}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'}\right]\Bigg|_{\mathbf{a}(\boldsymbol{\beta})=\mathbf{e}_t,\ \mathbf{c}(\boldsymbol{\beta})=\boldsymbol{\beta}} \\
&\overset{\text{yltä}}{=} \boldsymbol{\beta}'\mathbf{A}' \times \mathbf{0}_{m\times n} + \mathbf{e}_t^{'}\mathbf{A}\mathbf{I}_n \\
&= \mathbf{A}_t.
\end{aligned}$$

Kokoamalla tulokset t:n arvoille $t = 1, \ldots, m$ päällekkäin matriisimuotoon saadaan

$$\frac{\partial(\mathbf{A}\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'} = \begin{bmatrix} \dfrac{\partial(\mathbf{A}_1\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'} \\ \vdots \\ \dfrac{\partial(\mathbf{A}_m\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_m \end{bmatrix} = \mathbf{A}.$$

Nämä yhtälöt yksinäänkin todistavat väitteen ($\partial\mathbf{y}/\partial\mathbf{x}'$:n määritelmän mukaan).

Kaavan (1) mukaan $\partial\mathbf{y}'/\partial\mathbf{x} = [\partial\mathbf{y}/\partial\mathbf{x}']'$, joten

$$\frac{\partial(\boldsymbol{\beta}'\mathbf{A}')}{\partial\boldsymbol{\beta}} = \left[\frac{\partial(\mathbf{A}\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'}\right]' \overset{\text{yltä}}{=} [\mathbf{A}]' = \mathbf{A}'.$$

Vaihtoehtoinen johto b)-kohdan ensimmäiselle johdolle: Asetetaan a)-kohdassa $\mathbf{a}(\boldsymbol{\beta}) = 1$ (skalaari), $\mathbf{A}$:n paikalle $\mathbf{A}$:n t's vaakarivi eli $1 \times n$-vektori $\mathbf{A}_t$ $(t = 1,$

[1] Myöskään kaavassa (2) tulon $\mathbf{c}(\boldsymbol{\beta})'\mathbf{A}'\partial\mathbf{a}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}'$ termien dimensiot eivät täsmäisi. Ne olisivat $1 \times p$, $p \times n$ ja $1 \times m$ (eikä $n \times m$).

$\ldots, m$) ja $\mathbf{c}(\boldsymbol{\beta}) = \boldsymbol{\beta}$. Tällöin $\partial\mathbf{a}/\partial\boldsymbol{\beta}' = \mathbf{0}'$ $(1 \times n)$ ja $\partial\mathbf{c}/\partial\boldsymbol{\beta}' = \mathbf{I}_n$ ja

$$\begin{aligned}
\frac{\partial(\mathbf{A}_t\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'} &= \left[\frac{\partial\mathbf{a}(\boldsymbol{\beta})'\mathbf{A}_t\mathbf{c}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'}\right]\Bigg|_{\mathbf{a}(\boldsymbol{\beta})=1,\ \mathbf{c}(\boldsymbol{\beta})=\boldsymbol{\beta}} \\
&\overset{(2)\&\,\mathbf{A}\to\mathbf{A}'}{=} \left[\mathbf{c}(\boldsymbol{\beta})'\mathbf{A}_t'\frac{\partial\mathbf{a}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'} + \mathbf{a}(\boldsymbol{\beta})'\mathbf{A}_t\frac{\partial\mathbf{c}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'}\right]\Bigg|_{\mathbf{a}(\boldsymbol{\beta})=1,\ \mathbf{c}(\boldsymbol{\beta})=\boldsymbol{\beta}} \\
&= \boldsymbol{\beta}'\mathbf{A}_t \times \mathbf{0}' + 1 \times \mathbf{A}_t\mathbf{I}_n \\
&= \mathbf{A}_t.
\end{aligned}$$

Kuten yllä derivoimalla kullakin t:n arvolla $t = 1, \ldots, m$ ja kokoamalla tulokset matriisimuotoon saadaan

$$\frac{\partial(\mathbf{A}\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'} = \begin{bmatrix} \frac{\partial(\mathbf{A}_1\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'} \\ \vdots \\ \frac{\partial(\mathbf{A}_m\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_m \end{bmatrix} = \mathbf{A}.$$

c) Valitaan kaavassa (2) $n = m$ ja $\mathbf{a}(\boldsymbol{\beta}) = \mathbf{c}(\boldsymbol{\beta}) = \boldsymbol{\beta}$, jolloin $\partial\mathbf{a}/\partial\boldsymbol{\beta}' = \partial\mathbf{c}/\partial\boldsymbol{\beta}' = \mathbf{I}_m$:

$$\begin{aligned}
\frac{\partial(\boldsymbol{\beta}'\mathbf{A}\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'} &= \left[\frac{\partial\mathbf{a}(\boldsymbol{\beta})'\mathbf{A}\mathbf{c}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'}\right]\Bigg|_{\mathbf{a}(\boldsymbol{\beta})=\mathbf{c}(\boldsymbol{\beta})=\boldsymbol{\beta}} \\
&\overset{(2)}{=} \left[\mathbf{c}(\boldsymbol{\beta})'\mathbf{A}'\frac{\partial\mathbf{a}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'} + \mathbf{a}(\boldsymbol{\beta})'\mathbf{A}\frac{\partial\mathbf{c}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'}\right]\Bigg|_{\mathbf{a}(\boldsymbol{\beta})=\ \mathbf{c}(\boldsymbol{\beta})=\boldsymbol{\beta}} \\
&= \boldsymbol{\beta}'\mathbf{A}'\mathbf{I}_m + \boldsymbol{\beta}'\mathbf{A}\mathbf{I}_m \\
&= \boldsymbol{\beta}'(\mathbf{A}' + \mathbf{A}).
\end{aligned}$$

Kaavaa (1) hyväksi käyttäen saadaan edellisestä laskusta

$$\begin{aligned}
\frac{\partial(\boldsymbol{\beta}'\mathbf{A}\boldsymbol{\beta})}{\partial\boldsymbol{\beta}} &\overset{(1)}{=} \left(\frac{\partial(\boldsymbol{\beta}'\mathbf{A}\boldsymbol{\beta})}{\partial\boldsymbol{\beta}'}\right)' \\
&\overset{\text{ed.lasku}}{=} \{\boldsymbol{\beta}'(\mathbf{A}' + \mathbf{A})\}' \\
&= (\mathbf{A} + \mathbf{A}')\boldsymbol{\beta}.
\end{aligned}$$

Ensimmäisen yhtäsuuruus seuraa erityisen helposti huomaamalla, että $\boldsymbol{\beta}'\mathbf{A}\boldsymbol{\beta}$ on skalaari.

2.

$$\begin{aligned}
\frac{\partial \mathsf{S}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} &= \frac{\partial (\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \\
&= \frac{\partial}{\partial \boldsymbol{\beta}}(\mathbf{y}'\mathbf{y}-\mathbf{y}'\mathbf{X}\boldsymbol{\beta}-\boldsymbol{\beta}'\mathbf{X}'\mathbf{y}+\boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}) \\
&= \frac{\partial}{\partial \boldsymbol{\beta}}(\mathbf{y}'\mathbf{y}-\boldsymbol{\beta}'\mathbf{X}'\mathbf{y}-\boldsymbol{\beta}'\mathbf{X}'\mathbf{y}+\boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}) \\
&\overset{\text{1.1b\&symm.}}{=} -\mathbf{X}'\mathbf{y}-\mathbf{X}'\mathbf{y}+2\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \\
&= -2\mathbf{X}'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}).
\end{aligned}$$

Merkitsemällä gradienttivektori $\partial \mathsf{S}(\boldsymbol{\beta})/\partial \boldsymbol{\beta}$ nollavektoriksi saadaan normaaliyhtälöt

$$\mathbf{X}'(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\beta}}) = \mathbf{0} \tag{3}$$

($\hat{\boldsymbol{\beta}}$ on yhtälöt toteuttava vektori).

Lasketaan toisen kertaluvun ehto. Todetaan ensin, että gradienttivektori

$$\frac{\partial \mathsf{S}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = -2\mathbf{X}'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) = -2\mathbf{X}'\mathbf{y}+2\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \tag{4}$$

on k-vektori.Tehtävässä kehotetaan derivoimaan se vaakavektorin $\boldsymbol{\beta}'$ suhteen. Kaavan (4) oikeanpuolen ensimmäinen termi $-2\mathbf{X}'\mathbf{y}$ ei sisällä $\boldsymbol{\beta}$-vektorin komponentteja, joten derivoinnin tulos on $k \times k$ -nollamatriisi. Jälkimmäinen termi on HT 1.1 b)-kohdan muotoa, minkä perusteella saadaan

$$\frac{\partial 2\mathbf{X}'\mathbf{X}\boldsymbol{\beta}}{\partial \boldsymbol{\beta}'} = 2\mathbf{X}'\mathbf{X}$$

eli

$$\frac{\partial^2 \mathsf{S}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}'} = 2\mathbf{X}'\mathbf{X}.$$

Luennolla todistettiin, että $\mathbf{X}'\mathbf{X}$ on positiivisesti definiitti $k \times k$ -matriisi (kun $\mathbf{X}$ on täysiasteinen, mikä oletettiin tehtävässä), ja niin on tietenkin $2\mathbf{X}'\mathbf{X}$:kin. Näin ollen $[\partial^2 \mathsf{S}(\boldsymbol{\beta})/\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}']|_{\beta=\hat{\beta}}$ on positiivisesti definiitti matriisi eli neliösumma $\mathsf{S}(\boldsymbol{\beta})$ minimoituu (lokaalisti) $\boldsymbol{\beta}$:n arvolla $\hat{\beta}$.

Normaaliyhtälöistä seuraa, että

$$\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$$

eli että

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

3.[2] Tehtävänannon mukaan matriisi $\mathbf{A}$ $(m \times n)$ voidaan kirjoittaa

$$
\begin{aligned}
\mathbf{A} &= \begin{bmatrix} \mathbf{A}_{11} & \cdots & \mathbf{A}_{1p} \\ \vdots & & \vdots \\ \mathbf{A}_{q1} & \cdots & \mathbf{A}_{qp} \end{bmatrix} \\
&= \begin{bmatrix}
[\mathbf{A}_{11}]_{11} & & [\mathbf{A}_{11}]_{1n_1} & \cdots & [\mathbf{A}_{1p}]_{11} & \cdots & [\mathbf{A}_{1p}]_{1n_p} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{A}_{11}]_{m_1 1} & & [\mathbf{A}_{11}]_{m_1 n_1} & \cdots & [\mathbf{A}_{1p}]_{m_1 1} & \cdots & [\mathbf{A}_{1p}]_{m_1 n_p} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{A}_{q1}]_{11} & \cdots & [\mathbf{A}_{q1}]_{1n_1} & \cdots & [\mathbf{A}_{qp}]_{11} & \cdots & [\mathbf{A}_{qp}]_{1n_p} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{A}_{q1}]_{m_q 1} & \cdots & [\mathbf{A}_{q1}]_{m_q n_1} & \cdots & [\mathbf{A}_{qp}]_{m_q 1} & \cdots & [\mathbf{A}_{qp}]_{m_q n_p}
\end{bmatrix}.
\end{aligned}
$$

a) Koska matriisi $\mathbf{B}$ $(m \times n)$ on ositettu kuten $\mathbf{A}$, pätee

$$
\begin{aligned}
\mathbf{A} + \mathbf{B} &= \begin{bmatrix}
[\mathbf{A}_{11}]_{11} & & [\mathbf{A}_{11}]_{1n_1} & \cdots & [\mathbf{A}_{1p}]_{11} & \cdots & [\mathbf{A}_{1p}]_{1n_p} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{A}_{11}]_{m_1 1} & & [\mathbf{A}_{11}]_{m_1 n_1} & \cdots & [\mathbf{A}_{1p}]_{m_1 1} & \cdots & [\mathbf{A}_{1p}]_{m_1 n_p} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{A}_{q1}]_{11} & \cdots & [\mathbf{A}_{q1}]_{1n_1} & \cdots & [\mathbf{A}_{qp}]_{11} & \cdots & [\mathbf{A}_{qp}]_{1n_p} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{A}_{q1}]_{m_q 1} & \cdots & [\mathbf{A}_{q1}]_{m_q n_1} & \cdots & [\mathbf{A}_{qp}]_{m_q 1} & \cdots & [\mathbf{A}_{qp}]_{m_q n_p}
\end{bmatrix} \\
&+ \begin{bmatrix}
[\mathbf{B}_{11}]_{11} & & [\mathbf{B}_{11}]_{1n_1} & \cdots & [\mathbf{B}_{1p}]_{11} & \cdots & [\mathbf{B}_{1p}]_{1n_p} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{B}_{11}]_{m_1 1} & & [\mathbf{B}_{11}]_{m_1 n_1} & \cdots & [\mathbf{B}_{1p}]_{m_1 1} & \cdots & [\mathbf{B}_{1p}]_{m_1 n_p} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{B}_{q1}]_{11} & \cdots & [\mathbf{B}_{q1}]_{1n_1} & \cdots & [\mathbf{B}_{qp}]_{11} & \cdots & [\mathbf{B}_{qp}]_{1n_p} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{B}_{q1}]_{m_q 1} & \cdots & [\mathbf{B}_{q1}]_{m_q n_1} & \cdots & [\mathbf{B}_{qp}]_{m_q 1} & \cdots & [\mathbf{B}_{qp}]_{m_q n_p}
\end{bmatrix}
\end{aligned}
$$

[2]Vastaus on oleellisesti Jarkko Miettisen laatima.

$$
\begin{aligned}
&= \begin{bmatrix}
[\mathbf{A}_{11}]_{11}+[\mathbf{B}_{11}]_{11} & & [\mathbf{A}_{11}]_{1n_1}+[\mathbf{B}_{11}]_{1n_1} & \cdots & [\mathbf{A}_{1p}]_{11}+[\mathbf{B}_{1p}]_{11} & \cdots & [\mathbf{A}_{1p}]_{1n_p}+[\mathbf{B}_{1p}]_{1n_p} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{A}_{11}]_{m_1 1}+[\mathbf{B}_{11}]_{m_1 1} & & [\mathbf{A}_{11}]_{m_1 n_1}+[\mathbf{B}_{11}]_{m_1 n_1} & \cdots & [\mathbf{A}_{1p}]_{m_1 1}+[\mathbf{B}_{1p}]_{m_1 1} & \cdots & [\mathbf{A}_{1p}]_{m_1 n_p}+[\mathbf{B}_{1p}]_{m_1 n_p} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{A}_{q1}]_{11}+[\mathbf{B}_{q1}]_{11} & \cdots & [\mathbf{A}_{q1}]_{1n_1}+[\mathbf{B}_{q1}]_{1n_1} & \cdots & [\mathbf{A}_{qp}]_{11}+[\mathbf{B}_{qp}]_{11} & \cdots & [\mathbf{A}_{qp}]_{1n_p}+[\mathbf{B}_{qp}]_{1n_p} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{A}_{q1}]_{m_q 1}+[\mathbf{B}_{q1}]_{m_q 1} & \cdots & [\mathbf{A}_{q1}]_{m_q n_1}+[\mathbf{B}_{q1}]_{m_q n_1} & \cdots & [\mathbf{A}_{qp}]_{m_q 1}+[\mathbf{B}_{qp}]_{m_q 1} & \cdots & [\mathbf{A}_{qp}]_{m_q n_p}+[\mathbf{B}_{qp}]_{m_q n_p}
\end{bmatrix} \\
&= \begin{bmatrix}
\mathbf{A}_{11}+\mathbf{B}_{11} & \cdots & \mathbf{A}_{1p}+\mathbf{B}_{1p} \\
\vdots & & \vdots \\
\mathbf{A}_{q1}+\mathbf{B}_{q1} & \cdots & \mathbf{A}_{qp}+\mathbf{B}_{qp}
\end{bmatrix}.
\end{aligned}
$$

Käyttämällä transpoosin määritelmää nähdään, että

$$
\begin{aligned}
\mathbf{A}' &= \begin{bmatrix}
[\mathbf{A}_{11}]_{11} & & [\mathbf{A}_{11}]_{1n_1} & \cdots & [\mathbf{A}_{1p}]_{11} & \cdots & [\mathbf{A}_{1p}]_{1n_p} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{A}_{11}]_{m_1 1} & & [\mathbf{A}_{11}]_{m_1 n_1} & \cdots & [\mathbf{A}_{1p}]_{m_1 1} & \cdots & [\mathbf{A}_{1p}]_{m_1 n_p} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{A}_{q1}]_{11} & \cdots & [\mathbf{A}_{q1}]_{1n_1} & \cdots & [\mathbf{A}_{qp}]_{11} & \cdots & [\mathbf{A}_{qp}]_{1n_p} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{A}_{q1}]_{m_q 1} & \cdots & [\mathbf{A}_{q1}]_{m_q n_1} & \cdots & [\mathbf{A}_{qp}]_{m_q 1} & \cdots & [\mathbf{A}_{qp}]_{m_q n_p}
\end{bmatrix}' \\
&= \begin{bmatrix}
[\mathbf{A}_{11}]_{11} & & [\mathbf{A}_{11}]_{m_1 1} & \cdots & [\mathbf{A}_{q1}]_{11} & \cdots & [\mathbf{A}_{q1}]_{m_q 1} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{A}_{11}]_{1n_1} & & [\mathbf{A}_{11}]_{m_1 n_1} & \cdots & [\mathbf{A}_{q1}]_{1n_1} & \cdots & [\mathbf{A}_{q1}]_{m_q n_1} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{A}_{1p}]_{11} & \cdots & [\mathbf{A}_{1p}]_{m_1 1} & \cdots & [\mathbf{A}_{qp}]_{11} & \cdots & [\mathbf{A}_{qp}]_{m_q 1} \\
\vdots & & \vdots & & \vdots & & \vdots \\
[\mathbf{A}_{1p}]_{1n_p} & \cdots & [\mathbf{A}_{1p}]_{m_1 n_p} & \cdots & [\mathbf{A}_{qp}]_{1n_p} & \cdots & [\mathbf{A}_{qp}]_{m_q n_p}
\end{bmatrix} \\
&= \begin{bmatrix}
\mathbf{A}'_{11} & \cdots & \mathbf{A}'_{q1} \\
\vdots & & \vdots \\
\mathbf{A}'_{1p} & \cdots & \mathbf{A}'_{qp}
\end{bmatrix}.
\end{aligned}
$$

b) $\mathbf{AB}$ on $m \times r$ -matriisi. Merkitään matriisien $\mathbf{A}$ ja $\mathbf{B}$ i,j-elementtejä $[A]_{ij}$:llä ja $[B]_{lj}$:llä. Todistus alkaa matriisin $\mathbf{AB}$ mielivaltaisen i,j-elementin tarkastelulla ja siirtyy kolmannella rivillä koskemaan matriisin dimensioltaan $m_1 \times r_1$ olevan $1,1$-lohkon i,j-elementtiä (jolloin $1 \leq i \leq m_1$ ja $1 \leq j \leq r_1$):

$$\begin{aligned}
[\mathbf{AB}]_{ij} &= \textstyle\sum_{l=1}^{n} [A]_{il}[B]_{lj} \\
&= \textstyle\sum_{l=1}^{n_1} [A]_{il}[B]_{lj} + \sum_{l=n_1+1}^{n_1+n_2} [A]_{il}[B]_{lj} + ... + \sum_{l=n-n_p+1}^{n} [A]_{il}[B]_{lj} \\
&\overset{1,1\text{-lohko}}{\equiv} [\mathbf{A}_{11}\mathbf{B}_{11}]_{ij} + [\mathbf{A}_{12}\mathbf{B}_{21}]_{ij} + ... + [\mathbf{A}_{1p}\mathbf{B}_{p1}]_{ij} \\
&= \left[\textstyle\sum_{k=1}^{p} \mathbf{A}_{1k}\mathbf{B}_{k1}\right]_{ij}.
\end{aligned}$$

Matriisin $\mathbf{AB}$ $1,1$-lohko on siten $\sum_{k=1}^{p} \mathbf{A}_{ik}\mathbf{B}_{kj}$. Vastaava todistus pätee kaikille muillekin lohkoille (jolloin $m_{s-1} \leq i \leq m_s$ ja $r_{t-1} \leq j \leq r_t$, $s = 1, \ldots, q$, $t = 1, \ldots, p$ ja $m_0 \equiv r_0 \equiv 1$).

4. Algebra on yksinkertainen:

$$\mathbf{P_X X} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X} = \mathbf{X}.$$

Geometrista tulkintaa varten kirjoitetaan

$$\mathbf{P_X X} = \mathbf{P_X}\left[\mathbf{x}_1 \cdots \mathbf{x}_k\right] = \left[\mathbf{P_X x}_1 \cdots \mathbf{P_X x}_k\right] = \left[\mathbf{x}_1 \cdots \mathbf{x}_k\right].$$

Matriisi $\mathbf{P_X}$ projisoi mielivaltaisen (n-vektorin) $\mathbf{z}$ matriisin $\mathbf{X}$ sarakkeiden virittämään avaruuteen (kirjan s. 58). Matriisitulo $\mathbf{P_X X}$ voidaan ajatella ositettujen matriisien tuloiksi $[\mathbf{P_X x}_1 \cdots \mathbf{P_X x}_k]$. Koska jokainen $\mathbf{x}_j$-vektori, $j = 1, \ldots, n$, jo sijaitsee matriisin $\mathbf{X}$ sarakeavaruudessa, ei matriisilla $\mathbf{P_X}$ kertominen muuta niitä.

5. Olkoon $\mathbf{z}$ mielivaltainen k-vektori. Todetaan ensin, että $\underset{n\times k}{\mathbf{B}}\underset{k\times 1}{\mathbf{z}} \equiv [\mathbf{b}_1 \cdots \mathbf{b}_k]\,\mathbf{z} = \mathbf{0}$ jos ja vain jos $\mathbf{z} = \mathbf{0}$, sillä oletuksen mukaan $\mathbf{B}$ on täysiasteinen matriisi (sen sarakkeet ovat lineaarisesti riippumattomia). Oletetaan, että $\mathbf{z} \neq \mathbf{0}$, jolloin $\mathbf{Bz} \neq \mathbf{0}$. Tällöin $\mathbf{z}'\mathbf{B}'\mathbf{ABz} \overset{\mathbf{y}\equiv\mathbf{Bz}}{=} \mathbf{y}'\mathbf{Ay} > 0$, koska $\mathbf{A}$ on positiivisesti definiitti matrisiisi. Näin ollen myös $\mathbf{B}'\mathbf{AB}$ on positiivisesti definiitti matrisiisi.