

## University of Groningen

### Properties of the matrix A- XY

Steerneman, A.G.M.; van Perlo-ten Kleij, F.; Wong, A.

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2002

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Steerneman, A. G. M., van Perlo-ten Kleij, F., & Wong, A. (2002). *Properties of the matrix A- XY*. University of Groningen, SOM research school.

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# Properties of the matrix $A - XY'$

Ton Steerneman\*, Frederieke ten Kleij† and Amy Wong‡

October 10, 2002

**SOM-theme A Primary processes within firms**

## Abstract

The main topic of this paper is the matrix  $V = A - XY'$ , where  $A$  is a nonsingular  $k \times k$  matrix and  $X$  and  $Y$  are  $k \times p$  matrices of full column rank. Because properties of the matrix  $V$  can be derived from those of the matrix  $Q = I_k - XY'$ , we will consider in particular the case where  $A = I_k$ . For the case that  $Y'X = I_p$ , so that  $Q$  is singular, we will derive the Moore-Penrose inverse of  $Q$  in two ways. First, we generalize the result of Trenkler (2000) for  $p = 1$  and check whether this 'guess' satisfies the properties of the Moore-penrose inverse. Second, we will adopt a more elegant approach which exploits a decomposition of  $Q$  that is very similar to a singular value decomposition. An examination of the eigenvalues of  $Q$  leads to a decomposition that resembles an eigenvalue decomposition. Here we do not immediately impose that  $Y'X = I_p$ . Finally, we will focus on the eigenvalues and eigenvectors of the matrix  $D - xy'$ , with  $D$  diagonal.

---

The authors would like to thank Michel van de Velden for providing the reference Trenkler (2000).

\* Department of Econometrics, University of Groningen, P.O. Box 800, 9700 AV Groningen, THE NETHERLANDS. E-mail: a.g.m.steerneman@eco.rug.nl.

† Department of Econometrics, University of Groningen, P.O. Box 800, 9700 AV Groningen, THE NETHERLANDS. E-mail: f.ten.kleij@eco.rug.nl.

‡ Doctoral student Econometrics, University of Groningen, THE NETHERLANDS.

# 1 Introduction

In various (statistical) applications, we use a matrix of the type  $V = A - ab'$ , where  $A$  is some nonsingular  $k \times k$  matrix and  $a, b \in \mathbb{R}^k$ . The most well-known example is the centering operator with the matrix  $H = I_k - k^{-1}\iota_k\iota_k'$ , where  $\iota_k$  is a  $k \times 1$  vector of ones. This operator maps a vector  $(x_1, \dots, x_k)'$  to  $(x_1 - \bar{x}, \dots, x_k - \bar{x})'$ , where  $\bar{x}$  denotes the mean of the  $x_i$ . The matrix  $H$  is idempotent and it is the orthogonal projector on the hyperplane orthogonal to the vector  $\iota_k$ . Another well-known example is  $R = P - pp'$ , where  $p = (p_1, \dots, p_k)'$  with  $p_i > 0$ ,  $\iota_k'p = 1$ , and  $P = \text{diag}(p)$ . Note that  $nR$  is the covariance matrix of the multinomial distribution with parameters  $n$  and  $p$ . A third example is  $V = \text{diag}(e) + a\iota_k'$ , where  $e, a \in \mathbb{R}^k$  have positive elements. This matrix was originally studied by Vermeulen (1967), because of a physical investigation on the electronic properties of particle-counting diamonds.

The matrix  $V$ , with  $A$  symmetric, was studied in its general form by Trenkler (2000). He generalized results previously obtained by Vermeulen (1967), Klamkin (1970), Tanabe and Sagae (1992), Neudecker (1995), and Watson (1996). We will continue this line of research by, for example, dropping the assumption of symmetry of  $A$  and by replacing the vectors  $a$  and  $b$  by  $k \times p$  matrices, and also by considering the more special case  $V = D - ab'$ , where  $D = \text{diag}(d)$ .

In our notation, Vermeulen (1967) showed that the eigenvalues of  $D - a\iota_k'$  where  $D = \text{diag}(d)$  and  $a_1, \dots, a_k < 0$ , are real and positive. Klamkin (1970) gives a more elementary derivation. Moreover, he gives simple bounds for the eigenvalues. He derives the characteristic polynomial

$$|\lambda I_k - (D - a\iota_k')| = \left\{ 1 + \sum_{j=1}^k \frac{a_j}{\lambda - d_j} \right\} \prod_{i=1}^k (\lambda - d_i). \quad (1.1)$$

From this result, he observes that if  $0 < d_1 < d_2 < \dots < d_k$ , then the eigenvalues  $\lambda_1, \dots, \lambda_k$  are obtained by solving

$$1 + \sum_{j=1}^k \frac{a_j}{\lambda - d_j} = 0, \quad (1.2)$$

and if some of the  $d_i$  coincide, then there will be eigenvalues equal to the  $d_i$  that coincide. Trenkler (2000) generalizes this to  $D - ab'$ . In section 8, we will return to these results of Trenkler (2000) and give alternative proofs that exploit the original

ideas of Vermeulen (1967) and Klamkin (1970). Moreover, we will also derive the eigenvectors using some ideas of Watson (1996).

De Boer and Harkema (1984) were interested in the maximum likelihood estimation of sum constrained linear models:  $Y \sim \mathcal{N}_k(\mu, \Omega)$ ,  $t'_k Y = c$ , so that  $\Omega t_k = 0$ , where a certain structure will be imposed on  $\mu$ . Such models are of interest in modelling demand systems, brand choice, and so on. In case of relatively small samples, the model has to be parsimonious, especially with regard to the parameterization of  $\Omega$ . De Boer and Harkema (1984) suggested the specification

$$\Omega = D - \frac{1}{t'_k d} d d', \quad (1.3)$$

where  $D = \text{diag}(d)$  and  $d \in \mathbb{R}^k$ . Because of the constraint, they deleted one component of  $Y$  and the  $(k-1) \times (k-1)$  covariance matrix obtained became nonsingular. Wansbeek (1985) showed that estimation is possible without deletion of redundant observations. He assumed  $d_1 < \dots < d_k$  and obtained the following results. One eigenvalue of  $\Omega$  is equal to zero, the other eigenvalues satisfy

$$\sum_{i=1}^k \frac{d_i}{\lambda - d_i} = 0. \quad (1.4)$$

This follows from the characteristic equation he derived in the following way:

$$\begin{aligned} 0 = |\lambda I_k - \Omega| &= |\lambda I_k - D| \left( 1 + \frac{1}{t'_k d} d' (\lambda I_k - D)^{-1} d \right) \\ &= \frac{1}{t'_k d} |\lambda I_k - D| \left( t'_k (\lambda I_k - D) (\lambda I_k - D)^{-1} d + d' (\lambda I_k - D)^{-1} d \right) \\ &= |\lambda I_k - D| \left\{ \frac{\lambda}{t'_k d} t'_k (\lambda I_k - D)^{-1} d \right\}. \end{aligned}$$

Wansbeek (1985) observes from (1.4) that if  $d_i$  and  $d_{i+1}$  are of the same sign, then there lies an eigenvalue between them. We will use the same method in section 8 to obtain the characteristic polynomial of  $D - ab'$ . Wansbeek (1985) also gives the Moore-Penrose inverse of  $\Omega$ , namely  $\Omega^+ = H D^{-1} H$ . Since the matrix  $\Omega$  is symmetric and should be positive semi-definite, he concludes that  $0 < d_2 < \dots < d_k$  is a necessary condition. In case  $d_1 < 0 < d_2 < \dots < d_k$  he uses the Moore-Penrose inverse to establish that it is necessary that  $\sum_{i=1}^k d_i < 0$ . This can, however, more easily be shown by observing that it is necessary that the (1, 1) element of  $\Omega$  should be nonnegative. This amounts to

$$d_1 - \frac{1}{t'_k d} d_1^2 > 0.$$

Since  $d_1 < 0$ , we must have  $d_1/l'_k d > 1$ , hence,  $l'_k d < 0$ . In the present context it is, of course, more natural to require that all the  $d_i$  are positive.

A matrix that is very similar to  $\Omega$  is the matrix  $R$  we already discussed, since the covariance matrix of the multinomial distribution is based upon  $R = \text{diag}(p) - pp'$ , where  $p_1, \dots, p_k > 0$  and  $l'_k p \leq 1$ . If there are  $k + 1$  possible categories, then one may wish to count only the number of outcomes in the first  $k$  categories, because the number of outcomes in category  $k + 1$  uniquely follows from the total number of outcomes in the remaining categories. In this case  $l'_k p < 1$ . The matrix  $R$  has been studied under the condition  $l'_k p \leq 1$  by Tanabe and Sagae (1992). They obtained, among other things, the square-root free Cholesky decomposition, the Moore-Penrose inverse in case  $l'_k p = 1$ , namely  $R^+ = HP^{-1}H$ , and the inverse in case  $l'_k p < 1$ , that is,  $R^{-1} = P^{-1} + (1 - l'_k p)^{-1}l_k l'_k$ . Neudecker (1995) offered more elegant proofs and presents some new results. Watson (1996) assumes  $l'_k p = 1$  and shows how the eigenvalues and eigenvectors can be obtained. He shows that an eigenvalue not equal to any of the  $p_i$  should satisfy

$$\sum_{i=1}^k \frac{p_i^2}{p_i - \lambda} = 1. \quad (1.5)$$

This equation is very similar to (1.2) and (1.4). One eigenvalue is equal to zero and the other eigenvalues  $\lambda_1, \dots, \lambda_{k-1}$  satisfy

$$p_1 \leq \lambda_1 \leq p_2 \leq \lambda_2 \leq p_3 \leq \dots \leq \lambda_{k-1} \leq p_k$$

with strict inequalities if the  $p_i$ 's are all distinct. Similar observations are due to Klamkin (1970) and Wansbeek (1985). Watson furthermore derives how to obtain the eigenvectors. The product of the nonzero eigenvalues of  $R$  was obtained by Tanabe and Sagae (1992) and Neudecker (1995).

Dol (1991), and Dol, Steerneman and Wansbeek (1996) studied the Horvitz-Thompson estimator. Consider a finite population  $Y_1, \dots, Y_N$ . A fixed effective sample design of size  $n$  can be interpreted as a probability distribution on the set of all subsets of  $n$  elements from the labels  $\{1, \dots, N\}$ . Let  $S$  denote the random set of  $n$  labels that occur in the sample. The indicators  $E_1, \dots, E_N$  are defined by  $E_i = 1$  if  $i \in S$ , and  $E_i = 0$  if  $i \notin S$ . The first order inclusion probability is  $\pi_i = P(S \ni i)$  for  $i = 1, \dots, N$ . It is assumed that  $\pi_i = EE_i$  is positive. The Horvitz-Thompson estimator  $\bar{Y}_{HT}$  for the population mean  $\bar{Y}$  is

$$\bar{Y}_{HT} = \frac{1}{N} \sum_{i \in S} \frac{Y_i}{\pi_i} = \frac{1}{N} \sum_{i=1}^N E_i \frac{Y_i}{\pi_i}.$$

This a famous unbiased estimator. In order to give the variance, the second order inclusion probabilities are needed:  $\pi_{ij} = P(S \ni i, j) = EE_i E_j$ , for  $i, j = 1, \dots, N$ . Note that  $\pi_{ii} = \pi_i$ . We define  $\pi = (\pi_1, \dots, \pi_N)'$ ,  $\Pi = \text{diag}(\pi)$  and  $\Pi_2 = (\pi_{ij})$ . It is easy to see that  $\pi' \iota_N = n$  and  $\Pi_2 \iota_N = n\pi$ . The well-known expression for the variance of the Horvitz-Thompson estimator is

$$\text{Var } \bar{Y}_{HT} = N^{-2} Y' \Pi^{-1} (\Pi_2 - \pi \pi') \Pi^{-1} Y,$$

where  $Y = (Y_1, \dots, Y_N)'$ . The matrix  $\Pi_2 - \pi \pi'$  looks similar to  $R$ , but it is more complicated. In order to obtain bounds for this variance, Dol (1991), and Dol, Steerneman and Wansbeek (1996) obtained the following Moore-Penrose inverse:

$$(\Pi_2 - \pi \pi')^+ = H \Pi_2^{-1} H.$$

Inspired by Trenkler (2000), we will derive the (Moore-Penrose) inverse of

$$V = A - XY', \tag{1.6}$$

where  $A$  is a nonsingular  $k \times k$  matrix and  $X$  and  $Y$  are  $k \times p$  matrices of full column rank. Thus, we will generalize Trenkler's results in two ways. First, the matrix  $A$  is only restricted to be nonsingular, symmetry is not necessary. Secondly, the vectors  $a$  and  $b$  in the matrix  $A - ab'$  examined by Trenkler can be replaced by matrices of full column rank. Because  $|V| = |A| |I_p - Y'A^{-1}X|$ , an interesting case is  $Y'A^{-1}X = I_p$ , so that  $V$  is singular. We call this the singular case and it will be discussed in section 5. Note that the assumption that  $Y'A^{-1}X = I_p$  implies that both  $X$  and  $Y$  are of full-column rank. If  $|I_p - Y'A^{-1}X| \neq 0$ , then  $V$  is invertible, and we will refer to this as the nonsingular case. It will be discussed in section 4. We will not consider the mixture case where  $Y'A^{-1}X \neq I_p$  and  $|I_p - Y'A^{-1}X| = 0$ .

It is worthwhile to first consider a special case of (1.6), namely  $A = I_k$ , because properties of the matrix  $V$  can be derived from those of the matrix  $Q = I_k - XY'$ . If  $Y'X = I_p$ , then the matrix  $Q$  is idempotent, since  $Q^2 = Q$ . However, in general it is not symmetric. Some observations with regard to the rank of  $Q = I_k - XY'$ , which will prove their usefulness further on, are the following. For all vectors  $c$  with  $Y'c = 0$  we have that  $Qc = c$ , so that all vectors orthogonal to the columns of  $Y$  are eigenvectors of  $Q$  with eigenvalue 1. Because there exist  $(k - p)$  vectors in  $\mathbb{R}^k$  that are orthogonal to the  $p$  columns of  $Y$ , and since the rank of a square matrix equals the number of nonzero eigenvalues, we know that  $\text{rank}(Q) \geq k - p$ . If, in addition,  $Y'X = I_p$ , we know that all eigenvalues of  $Q$  are all equal to 0 or 1, since in this case  $Q$  is idempotent. Because in this case  $QX = 0$ , the eigenvectors corresponding to  $\lambda = 0$  are the  $p$  columns of the matrix  $X$ . Therefore, we know that  $Q$  has at least  $p$  eigenvalues equal to zero, so that

$\text{rank}(Q) \leq k - p$ . It now immediately follows that in the special case where  $Y'X = I_p$ ,  $\text{rank}(Q) = k - p$ .

An alternative way to determine the rank of  $Q$  is to look at its characteristic equation:

$$\begin{aligned}
|\lambda I_k - Q| &= |(\lambda - 1)I_k + XY'| \\
&= (\lambda - 1)^k |I_k + \frac{1}{\lambda - 1}XY'| \\
&= (\lambda - 1)^k |I_p + \frac{1}{\lambda - 1}Y'X| \\
&= (\lambda - 1)^{k-p} |(\lambda - 1)I_p + Y'X|, \tag{1.7}
\end{aligned}$$

for  $\lambda \neq 1$ . From (1.7) we observe that  $Q$  has at least  $k - p$  eigenvalues equal to 1, as we already noticed above. In the special case that  $Y'X = I_p$ , equation (1.7) simplifies to

$$|\lambda I_k - Q| = (\lambda)^p (\lambda - 1)^{k-p}. \tag{1.8}$$

Therefore, if  $Y'X = I_p$ , the eigenvalue  $\lambda = 0$  has multiplicity  $p$  and the other  $k - p$  eigenvalues are equal to 1. This once more shows that  $\text{rank}(Q) = k - p$ .

In particular, if  $V = A - XY'$  with  $Y'A^{-1}X = I_p$ , it is not difficult to see that the rank of  $V$  also equals  $k - p$ , since  $V = A - XY' = A(I_k - A^{-1}XY')$ , where  $Y'A^{-1}X = I_p$ ,  $\text{rank}(V) = \text{rank}(I_k - A^{-1}XY')$ , because  $A$  is nonsingular. However, the rank of  $(I_k - A^{-1}XY')$  equals  $k - p$  as we showed above. Because  $VA^{-1}X = 0$ , the eigenvectors corresponding to  $\lambda = 0$  are the  $p$  columns of the matrix  $A^{-1}X$ .

In section 2 we present basic notation on generalized inverses and in section 3 some general results on idempotent matrices. Subsequently, we will shortly address the case where  $V$  is nonsingular in section 4. Section 5 deals with the singular case, where we will first consider the specific situation where  $A = I_k$ . The general case then easily follows. In section 6 we obtain a kind of singular value decomposition for  $Q$  if  $Y'X = I_p$ . An examination of the eigenvalues of  $Q$  in section 7 leads to a decomposition that resembles an eigenvalue decomposition. Here we do not immediately impose that  $Y'X = I_p$ . Finally, section 8 focuses on the eigenvalues and eigenvectors of the matrix  $D - xy'$ , with  $D$  diagonal.

## 2 The Moore-Penrose inverse: some preliminaries

Let  $A$  be a  $k \times p$  matrix and consider the  $p \times k$  matrix  $X$  which satisfies one or more of the following properties:

- (1)  $AXA = A$ ,
- (2)  $XAX = X$ ,
- (3)  $XA$  is symmetric,
- (4)  $AX$  is symmetric.

If  $X$  satisfies (1), then  $X$  is called a generalized inverse of  $A$ , denoted by  $X = A^-$ . If  $X$  satisfies both (1) and (2), then  $X$  is called a reflexive generalized inverse of  $A$ , which is denoted by  $X = A_r^-$ . If  $X$  satisfies the properties (1), (2) and (3), then we call  $X$  a left pseudoinverse of  $A$ , denoted by  $A_L^-$ , whereas we call  $X$  a right pseudoinverse of  $A$ , denoted by  $A_R^-$ , if it satisfies the properties (1), (2) and (4). Finally, if  $X$  satisfies all four properties, then  $X$  is called the Moore-Penrose inverse of  $A$  which we will denote by  $A^+$ . The Moore-Penrose inverse of a matrix is uniquely defined by (1)–(4). For textbooks on generalized inverses we refer to, for example, Bouillion and Odell (1971) and Rao and Mitra (1971).

**Lemma 1.** *The matrix  $A_L^- A A_R^-$  is the Moore-Penrose inverse of  $A$ .*

This lemma is easily proved by checking the four conditions the Moore-Penrose inverse has to satisfy (Bouillion and Odell, 1971, chapter 1).

## 3 Properties of idempotent matrices

As already mentioned in section 1, the matrix  $Q = I_k - XY'$  is idempotent if  $Y'X = I_p$ . A typical example is  $Y' = X^+ = (X'X)^{-1}X'$ . In this case  $Q = I_k - XX^+ = I_k - X(X'X)^{-1}X'$  is the very familiar symmetric, idempotent matrix to be denoted by  $P_X$ : the orthogonal projector on the orthogonal complement of the column space of  $X$ . To give another example in which this type of matrix appears, but now as an orthogonal projector with respect to another inner product, consider the standard linear regression model

$$y = X\beta + \varepsilon,$$

where  $\varepsilon \sim \mathcal{N}_k(0, \sigma^2 \Sigma)$ , and  $\Sigma$  is known. According to the method of Generalized Least Squares (GLS), we have to minimize  $(y - X\beta)' \Sigma^{-1} (y - X\beta)$ . Here, the underlying inner product is  $(a, b) = a' \Sigma^{-1} b$ . The solution is  $\hat{\beta} = (X' \Sigma^{-1} X)^{-1} X' \Sigma^{-1} y$ , the Aitken estimator, so the GLS approximation to  $y$  is  $X(X' \Sigma^{-1} X)^{-1} X' \Sigma^{-1} y$ . This is the



orthogonal projection of  $y$  on  $\{X\beta | \beta \in \mathbb{R}^p\}$  with respect to  $\Sigma^{-1}$ . The vector of residuals is

$$e = (I_k - X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1})y.$$

This leads to

$$\begin{aligned} Q &= I_k - X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1} \\ &= I_k - XY', \end{aligned}$$

where  $Y = \Sigma^{-1}X(X'\Sigma^{-1}X)^{-1}$  and  $Y'X = I_p$ . Conversely, such a matrix  $Q$  can be interpreted as an orthogonal projector. Given  $Q = I_k - XY'$  with  $Y'X = I_p$ , we take, for example,  $\Sigma^{-1} = YY' + P_X$ . This matrix is symmetric and positive definite. To establish this last property, note that  $x'\Sigma^{-1}x = 0$  is equivalent to  $Y'x = 0$  and  $x = X\beta$  for some  $\beta \in \mathbb{R}^p$ . So,  $0 = Y'x = Y'X\beta = \beta \Rightarrow x = 0$ .

Since we also use the properties of idempotent matrices, we mention the most important facts. A  $k \times k$  matrix  $Q$  is idempotent if  $Q^2 = Q$ . In statistics and econometrics,  $Q$  will often also be symmetric, but this is not necessary as we have remarked. If  $Q$  is idempotent, then  $I_k - Q$  is also idempotent and  $Q(I_k - Q) = (I_k - Q)Q = 0$ .

An important idempotent matrix in this paper is the matrix  $P_U = I_k - U(U'U)^{-1}U' = I_k - UU^+$ , where  $U$  is a  $k \times p$  matrix of full column rank. Because  $P_U$  is also symmetric, it is a projection matrix.  $P_U$  is in fact the orthogonal projector on the orthogonal complement of the column space of  $U$ , so that  $P_U U = 0$  and  $(I_k - P_U)U = U$ . If  $V$  is another  $k \times p$  matrix with columns orthogonal to the columns of  $U$ , then  $P_U V = V$ . Moreover, in this case the corresponding projection matrices of  $U$  and  $V$  commute, that is, if  $U'V = 0$ ,

$$P_U P_V = P_V P_U = P_{(U \ V)}. \quad (3.1)$$

Because the matrix  $Q = I_k - XY'$  plays a key role, we will now discuss some properties of this matrix that will facilitate derivations further on. We assume that the  $k \times p$  matrices  $X$  and  $Y$  are of rank  $p$  and that  $Y'X = I_p$ . Let  $R(A)$  denote the linear subspace spanned by the columns of the matrix  $A$ . We can distinguish three cases: (i) the columns of  $X$  and  $Y$  span different spaces, that is,  $R(X) \cap R(Y) = \{0\}$ , so that  $\text{rank}(X, Y) = 2p$ , (ii)  $\text{rank}(X, Y) = p$ , which means that  $X$  and  $Y$  span the same space, and (iii)  $p < \text{rank}(X, Y) < 2p$ , so that there is partial overlap between the column spaces of  $X$  and  $Y$ . The first case will be discussed extensively in section 6 and 7.

If  $X$  and  $Y$  span the same space, case (ii), then there exists a nonsingular  $k \times k$  matrix  $B$  such that  $X = YB$ . Therefore,  $Y'X = Y'YB = I_p$ , so that  $B = (Y'Y)^{-1}$  and

$Q = I_k - Y(Y'Y)^{-1}Y' = P_Y$ . So, in this specific situation  $Q = P_X = P_Y$  is symmetric and idempotent. It immediately follows that  $Q^+ = P_X = P_Y$ .

If there is overlap between the column spaces of  $X$  and  $Y$ , we assume that  $X = (X_1, X_2)$ ,  $Y = (Y_1, Y_2)$ , where  $X_1$  and  $Y_1$  are  $k \times q$  matrices,  $X_2$  and  $Y_2$  are  $k \times r$  matrices with  $q + r = p$ , such that  $R(X_1) = R(Y_1)$  and  $R(X_2) \cap R(Y_2) = \{0\}$ . Note that we can write  $Q$  as  $Q = I_k - X_1Y_1' - X_2Y_2'$ . From the identity  $Y'X = I_p$  we obtain the properties

$$Y_1'X_1 = I_q, \quad Y_2'X_2 = I_r \quad (3.2)$$

$$Y_1'X_2 = 0, \quad Y_2'X_1 = 0 \quad (3.3)$$

Because  $X_1$  and  $Y_1$  span the same space, the same argument as above applies, so that we can write  $Y_1 = X_1(X_1'X_1)^{-1}$  and  $X_1 = Y_1(Y_1'Y_1)^{-1}$ . Because  $Y_1'X_2 = (X_1'X_1)^{-1}X_1'X_2 = 0$ , it follows that  $X_1'X_2 = 0$ , and we see that the columns of  $X_1$  are mutually orthogonal to those of  $X_2$ . Analogously, it can be shown that  $Y_2'Y_1 = 0$ . These observations lead to the following lemma:

**Lemma 2.** *Let  $Q = I_k - XY'$ , where  $X = (X_1, X_2)$ ,  $Y = (Y_1, Y_2)$ ,  $X_1$  and  $Y_1$  are  $k \times q$  matrices,  $X_2$  and  $Y_2$  are  $k \times r$  matrices with  $q + r = p$ , such that  $R(X_1) = R(Y_1)$  and  $R(X_2) \cap R(Y_2) = \{0\}$ . Define  $Q_2 = I_k - X_2Y_2'$ . Then  $Q$  can be written as the product of two idempotent matrices that commute:*

$$Q = P_{X_1}Q_2 = Q_2P_{X_1}. \quad (3.4)$$

*Proof.* To establish (3.4), observe that

$$\begin{aligned} P_{X_1}Q_2 &= P_{X_1} - P_{X_1}X_2Y_2' = P_{X_1} - X_2Y_2' = Q \\ Q_2P_{X_1} &= Q_2 - Q_2X_1Y_1' = Q_2 - X_1Y_1' + X_2Y_2'X_1Y_1' = Q. \end{aligned}$$

□

## 4 The nonsingular case

It is well-known that the matrix

$$S = \begin{pmatrix} A & X \\ Y' & I_p \end{pmatrix}$$

can be written as

$$S = \begin{pmatrix} I_k & 0 \\ Y'A^{-1} & I_p \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & I_p - Y'A^{-1}X \end{pmatrix} \begin{pmatrix} I_k & A^{-1}X \\ 0 & I_p \end{pmatrix}. \quad (4.1)$$

This representation is very instructive, since it immediately follows that  $S$  is nonsingular if and only if  $I_p - Y'A^{-1}X$  is nonsingular. Moreover, equation (4.1) also shows that  $|S| = |A||I_p - Y'A^{-1}X|$ . If  $S$  is nonsingular, we know from the standard results on inverses of partitioned matrices that  $S^{11}$ , the upper left-hand block of  $S^{-1}$ , can be written in two ways:

$$\begin{aligned} S^{11} &= A^{-1} + A^{-1}X(I_p - Y'A^{-1}X)^{-1}Y'A^{-1} & (4.2) \\ &= (A - XY')^{-1}, & (4.3) \end{aligned}$$

and we have a well-known expression for  $(A - XY')^{-1}$  (see e.g. Rao and Mitra, 1971, chapter 2). If  $S$  is singular, it is tempting to replace the inverses by Moore-Penrose inverses. According to corollary 4.4 from Ouellette (1981), we have the following result.

**Theorem 1.** *If  $A$  is a  $k \times k$  matrix,  $X$  and  $Y$  are  $k \times p$  matrices with  $k \geq p$ , and if*

$$\text{rank} \begin{pmatrix} A \\ Y' \end{pmatrix} = \text{rank}(A, X) = \text{rank } A = \text{rank}(A - XY')$$

and

$$\text{rank}(I_p - Y'A^+X) = p,$$

then

$$(A - XY')^+ = A^+ + A^+X(I_p - Y'A^+X)^{-1}Y'A^+. \quad (4.4)$$

On account of theorem 4.6 from Ouellette (1981), which originates from Marsaglia and Styan (1974), page 439, we know that we need  $I_p - Y'A^+X$  to be nonsingular in order to have results similar to (4.2) and (4.4).

However, in the sequel, we will focus on the case that  $I_p = Y'A^+X$ . In particular, we already assumed that  $A$  is nonsingular. Note that theorem 1 also does not apply to the mixture case where  $Y'A^{-1}X \neq I_p$  and  $|I_p - Y'A^{-1}X| = 0$ , as mentioned in section 1.

## 5 The singular case

In this section we will be interested in obtaining the Moore-Penrose inverse of  $V = A - XY'$ , where the  $k \times k$  matrix  $A$  is nonsingular and the  $k \times p$  matrices  $X$  and  $Y$  satisfy the condition  $Y'A^{-1}X = I_p$ . As we observed, this implies that  $X$  and  $Y$  are of full column rank  $p \leq k$ . Inspired by Trenkler (2000) the following result can be guessed. We will show that it is indeed correct.

**Theorem 2.** Let  $V = A - XY'$ , where  $A$  is a nonsingular  $k \times k$  matrix, and  $X$  and  $Y$  are  $k \times p$  matrices with  $Y'A^{-1}X = I_p$ . Then

$$V^+ = P_H A^{-1} P_K,$$

where  $H = A^{-1}X$  and  $K = (A^{-1})'Y$ , is the Moore-Penrose inverse of  $V$ .

The theorem can be established by verifying the four conditions for the Moore-Penrose inverse. However, we think that it is nicer to obtain the result in the special case that  $A = I_k$  first, and then to derive the more general result in a constructive way.

We first focus on  $Q = I_k - XY'$  where  $X$  and  $Y$  are  $k \times p$  matrices with  $Y'X = I_p$ . Some useful properties are:

$$QX = 0, \quad Y'Q = 0 \quad (5.1)$$

$$QP_X = Q, \quad P_X Q = P_X \quad (5.2)$$

$$P_Y Q = Q, \quad QP_Y = P_Y \quad (5.3)$$

$$QY = -P_Y XY'Y \quad X'Q = -X'XY'P_X \quad (5.4)$$

$$QQ = Q \quad (5.5)$$

From (5.5) we see that  $Q$  is idempotent, but not necessarily symmetric. Later on, we will give a decomposition of  $Q$  that is very similar to a singular value decomposition. From this result  $Q^+$  can be derived in a constructive way, see section 6. Checking the four conditions, however, is easier. Obviously, we have from (5.2), (5.3) and (5.5) that  $QP_X P_Y Q = QQ = Q$ , so that  $P_X P_Y$  is a generalized inverse of  $Q$ . Next, we observe that  $P_Y Q P_X = Q$ . Hence  $P_X P_Y Q P_X P_Y = P_X Q P_Y = P_X P_Y$ . Moreover,  $P_X P_Y Q = P_X Q = P_X$  and  $Q P_X P_Y = P_Y$  are symmetric matrices. These observations prove the following theorem.

**Theorem 3.** Let  $Q = I_k - XY'$ , where  $X$  and  $Y$  are  $k \times p$  matrices with  $Y'X = I_p$ . Then

$$Q^+ = P_X P_Y.$$

The proof of theorem 2 can now easily be obtained from theorem 3 by applying lemma 1.

*Proof of theorem 2.* We will derive the Moore-Penrose inverse by using a left and right pseudoinverse of  $V$ , cf. lemma 1. Note that

$$\begin{aligned} A - XY' &= A(I_k - A^{-1}XY') \\ &= (I_k - XY'A^{-1})A. \end{aligned}$$

This suggests to consider  $(I_k - A^{-1}XY')^+A^{-1}$  and  $A^{-1}(I_k - XY'A^{-1})^+$ , to be denoted by  $V_L^-$  and  $V_R^-$  respectively. Obviously,  $V_L^-$  is indeed a left pseudoinverse of  $V$  and  $V_R^-$  is a right pseudoinverse of  $V$ . Lemma 1 states that the Moore-Penrose of  $V$  can now be computed as  $V^+ = V_L^- V V_R^-$ . From theorem 3 we know that  $(I_k - A^{-1}XY')^+ = P_H P_Y$  and  $(I_k - XY'A^{-1})^+ = P_X P_K$ . On account of (5.2), it follows that

$$\begin{aligned} V^+ &= P_H P_Y A^{-1} (A - XY') A^{-1} P_X P_K \\ &= P_H (I_k - HY') A^{-1} P_K \\ &= P_H (A^{-1} - XK') P_K \\ &= P_H A^{-1} P_K. \end{aligned}$$

□

Taking  $X = a$  and  $Y = -b$ , we have the result derived in Trenkler (2000). If we compare the expression of Trenkler for the Moore-Penrose inverse of  $A + ab'$  with  $V^+$ , then we see that our result is a straightforward generalization. We therefore could have guessed this solution and verify the four conditions the Moore-Penrose inverse has to satisfy, just as we did in the proof of theorem 3. Anyway, the basic properties (5.1)–(5.5) of idempotent matrices like  $Q$  are needed. We think, however, that the proof as given above is nicer. Straightforward multiplication shows that

$$V V^+ = I_k - K K^+ \quad (5.6)$$

$$V^+ V = I_k - H H^+, \quad (5.7)$$

where we used the fact that  $K^+ = (K'K)^{-1}K'$ , because  $K$  is of full-column rank. Thus,  $V V^+$  and  $V^+ V$  are symmetric. From (5.6) and (5.7) it now easily follows that  $V V^+ V = V$  and  $V^+ V V^+ = V^+$ , so that indeed all four conditions hold.

In section 3, we distinguished three cases with respect to the spaces spanned by the columns of  $X$  and  $Y$ . Although the Moore-Penrose inverse of  $Q$  is given by  $P_X P_Y$ , regardless of the relation between  $R(X)$  and  $R(Y)$ , it can also be found by exploiting the specific structure of  $Q$  in these three cases. Case (i), where  $R(X) \cap R(Y) = \{0\}$ , so that  $\text{rank}(X, Y) = 2p$ , will be discussed extensively in section 6. If  $R(X) = R(Y)$ , case (ii), we observed that  $Q = P_X = P_Y$ , so that  $Q^+ = P_X = P_Y$ . For case (iii), where  $p < \text{rank}(X, Y) < 2p$ , we know from lemma 2 that  $Q = P_{X_1} Q_2 = Q_2 P_{X_1}$ . A natural guess for the Moore-Penrose inverse of  $Q$  is now:

$$Q^+ = P_{X_1} Q_2^+ = Q_2^+ P_{X_1}. \quad (5.8)$$

From (3.1) and section 3 we know that  $P_{X_1} = P_{Y_1}$  and that  $P_{X_1}$  and  $P_{X_2}$  respectively  $P_{Y_1}$  and  $P_{Y_2}$  commute. From theorem 3 we know that

$$Q^+ = P_{X_1} P_{X_2} P_{Y_1} P_{Y_2} = P_{X_1} P_{X_2} P_{X_1} P_{Y_2} = P_{X_1} P_{X_2} P_{Y_2} = P_{X_1} Q_2^+.$$

On the other hand, we have

$$P_{X_1} Q_2^+ = P_{X_1} P_{X_2} P_{Y_2} = P_{X_2} P_{Y_2} P_{X_1} = Q_2^+ P_{X_1}.$$

## 6 A blockwise singular value decomposition

In this section we will present a decomposition of  $Q = I_k - XY'$  which is quite similar to a singular value decomposition. It is assumed that  $X$  and  $Y$  are  $k \times p$  matrices of rank  $p$  and  $Y'X = I_p$ . We will only consider the case that the columns of  $X$  and  $Y$  span different spaces, that is,  $R(X) \cap R(Y) = \{0\}$ , so that  $\text{rank}(X, Y) = 2p$ . Analogous to a singular value decomposition, we are looking for orthogonal matrices  $S$  and  $T$  and a matrix  $\Lambda$  such that  $Q = S\Lambda T'$ . As opposed to the singular value decomposition, however, we do not restrict  $\Lambda$  to be strictly diagonal, although an easy structure is indeed convenient. We will take  $\Lambda$  to be block-diagonal. This decomposition provides us an alternative method to find the Moore-Penrose inverse of  $Q$ , because it can be easily checked that in this case  $Q^+ = T\Lambda^+ S'$  also holds.

Because  $QX = 0$  and  $Q'Y = 0$ , we know that the columns of  $X$  and  $Y$  are right and left singular vectors of  $Q$  with singular value 0. Moreover, if the  $k \times 1$  vector  $a$  is orthogonal to the columns of  $X$  and  $Y$ , then  $Qa = a$  and  $Q'a = a$ . So  $a$  is both a left and a right singular vector of  $Q$  with singular value 1. Since the columns of  $X$  and  $Y$  constitute an independent system of  $2p$  vectors in  $\mathbb{R}^k$ , we can find vectors  $w_1, \dots, w_{k-2p}$  that are mutually orthogonal, have unit length and are orthogonal to the columns of  $X$  and  $Y$ . If  $W = (w_1, \dots, w_{k-2p})$ , then  $QW = W$  and  $Q'W = W$ .

We are looking for a decomposition  $Q = S\Lambda T'$  where  $S$  and  $T$  are orthogonal  $k \times k$  matrices. We would like that  $S$  is composed mainly of left singular vectors and  $T$  of right singular vectors. Let  $S = (S_1, S_2, S_3)$ ,  $T = (T_1, T_2, T_3)$  and

$$\Lambda = \begin{pmatrix} \Lambda_1 & 0 & 0 \\ 0 & \Lambda_2 & 0 \\ 0 & 0 & \Lambda_3 \end{pmatrix},$$

then obviously we can take  $S_3 = T_3 = W$  and  $\Lambda_3 = I_{k-2p}$ . We observe that  $S_1$  should then be build up from columns of  $Y$ ,  $T_1$  should accordingly be constructed from  $X$ , and we take  $\Lambda_1 = 0_p$ . This leads to the choice  $S_1 = Y(Y'Y)^{-\frac{1}{2}} \equiv \tilde{Y}$  and  $T_1 = X(X'X)^{-\frac{1}{2}} \equiv \tilde{X}$ , because then  $S_1' S_1 = T_1' T_1 = I_p$ ,  $S_1' S_3 = 0$  and  $T_1' T_3 = 0$ . The columns of  $S_2$  should be orthogonal to  $S_1$  and  $S_3$  and have to be constructed from  $\tilde{X}$ , because the columns of  $\tilde{X}$ ,  $\tilde{Y}$  and  $W$  provide a basis in  $\mathbb{R}^k$ . This leads to the choice of  $P_Y \tilde{X}$  and the orthonormal version is

$$S_2 = P_Y \tilde{X} (\tilde{X}' P_Y \tilde{X})^{-\frac{1}{2}} = P_Y \tilde{X} H_{X'Y}^{-\frac{1}{2}},$$

with  $H_{XY} \equiv \tilde{X}' P_Y \tilde{X}$ . Analogously, we use

$$T_2 = -P_X \tilde{Y} (\tilde{Y}' P_X \tilde{Y})^{-\frac{1}{2}} = -P_X \tilde{Y} H_{Y\tilde{X}}^{-\frac{1}{2}},$$

with  $H_{Y\tilde{X}} \equiv \tilde{Y}' P_X \tilde{Y}$ . It turned out that we need a minus sign for  $T_2$ . In order to obtain  $\Lambda_2$  we solve  $QT_2 = S_2 \Lambda_2$  using (5.4):

$$\begin{aligned} QT_2 &= -Q P_X \tilde{Y} H_{Y\tilde{X}}^{-\frac{1}{2}} \\ &= -Q \tilde{Y} H_{Y\tilde{X}}^{-\frac{1}{2}} \\ &= P_Y X (Y' Y)^{\frac{1}{2}} H_{Y\tilde{X}}^{-\frac{1}{2}} \\ &= S_2 \Lambda_2, \end{aligned}$$

where

$$\begin{aligned} \Lambda_2 &= H_{\tilde{X}Y}^{\frac{1}{2}} (X' X)^{\frac{1}{2}} (Y' Y)^{\frac{1}{2}} H_{Y\tilde{X}}^{-\frac{1}{2}} \\ &= H_{\tilde{X}Y}^{\frac{1}{2}} (\tilde{Y}' \tilde{X})^{-1} H_{Y\tilde{X}}^{-\frac{1}{2}}. \end{aligned} \quad (6.1)$$

**Theorem 4.** Let  $X, Y$  be  $k \times p$  matrices of rank  $p$ , such that  $Y' X = I_p$ ,  $R(X) \cap R(Y) = \{0\}$ . Define  $\tilde{X} = X(X' X)^{-\frac{1}{2}}$ ,  $\tilde{Y} = Y(Y' Y)^{-\frac{1}{2}}$ ,  $H_{XY} = \tilde{X}' P_Y \tilde{X}$  and  $H_{Y\tilde{X}} = \tilde{Y}' P_X \tilde{Y}$ . Then  $Q = I_k - XY'$  can be decomposed as  $Q = S \Lambda T'$ , where

$$\begin{aligned} \Lambda &= \text{diag}(0_p, H_{\tilde{X}Y}^{\frac{1}{2}} (\tilde{Y}' \tilde{X})^{-1} H_{Y\tilde{X}}^{-\frac{1}{2}}, I_{k-2p}) \\ S &= (\tilde{Y}, P_Y \tilde{X} H_{\tilde{X}Y}^{-\frac{1}{2}}, W) \\ T &= (\tilde{X}, -P_X \tilde{Y} H_{Y\tilde{X}}^{-\frac{1}{2}}, W), \end{aligned}$$

such that  $S$  and  $T$  are orthogonal matrices. The  $k \times (k - 2p)$  matrix  $W$  has the property that  $W' W = I_{k-2p}$  and  $Y' W = X' W = 0$ .

The following properties are easily derived and will be useful in further derivations.

$$\tilde{X}' \tilde{X} = I_p \quad \tilde{Y}' \tilde{Y} = I_p \quad (6.2)$$

$$\tilde{X}' \tilde{Y} = (X' X)^{-\frac{1}{2}} (Y' Y)^{-\frac{1}{2}} \quad \tilde{Y}' \tilde{X} = (Y' Y)^{-\frac{1}{2}} (X' X)^{-\frac{1}{2}} \quad (6.3)$$

$$P_X = I_k - \tilde{X} \tilde{X}' \quad P_Y = I_k - \tilde{Y} \tilde{Y}' \quad (6.4)$$

$$H_{XY} = I_k - \tilde{X}' \tilde{Y} \tilde{Y}' \tilde{X} \quad H_{Y\tilde{X}} = I_k - \tilde{Y}' \tilde{X} \tilde{X}' \tilde{Y} \quad (6.5)$$

If  $p = 1$ , the equations (6.3) and (6.5) have the following interpretation:

$$\begin{aligned}\tilde{X}'\tilde{Y} &= \tilde{Y}'\tilde{X} = \cos\theta \\ H_{XY} &= H_{YX} = \sin^2\theta,\end{aligned}$$

where  $\theta$  denotes the angle between  $X$  and  $Y$ . From (6.2)-(6.5) it follows that

$$\tilde{Y}'\tilde{X}H_{XY} = H_{YX}\tilde{Y}'\tilde{X},$$

so that

$$\tilde{Y}'\tilde{X}H_{XY}^{-1} = H_{YX}^{-1}\tilde{Y}'\tilde{X}. \quad (6.6)$$

Now we can compute the Moore-Penrose inverse of  $Q$  as  $Q^+ = T\Lambda^+S' = T_2\Lambda_2^{-1}S_2' + WW'$ , where

$$\Lambda_2^{-1} = H_{YX}^{\frac{1}{2}}\tilde{Y}'\tilde{X}H_{XY}^{-\frac{1}{2}},$$

and  $WW' = P_X - P_X\tilde{Y}H_{YX}^{-1}\tilde{Y}'P_X$ , because  $TT' = T_1T_1' + T_2T_2' + WW' = I_k$ . Using (6.2)-(6.6), we obtain

$$\begin{aligned}Q^+ &= -P_X\tilde{Y}(\tilde{Y}'\tilde{X})H_{XY}^{-1}\tilde{X}'P_Y + P_X - P_X\tilde{Y}H_{YX}^{-1}\tilde{Y}'P_X \\ &= P_X\left[I_k - \tilde{Y}H_{YX}^{-1}\tilde{Y}'P_X - \tilde{Y}H_{YX}^{-1}\tilde{Y}'\tilde{X}\tilde{X}'P_Y\right] \\ &= P_X\left[I_k - \tilde{Y}H_{YX}^{-1}\tilde{Y}'(P_X + \tilde{X}\tilde{X}'P_Y)\right] \\ &= P_X\left[I_k - \tilde{Y}H_{YX}^{-1}\tilde{Y}'(I_k - \tilde{X}\tilde{X}'\tilde{Y}\tilde{Y}')\right] \\ &= P_X\left[I_k - \tilde{Y}H_{YX}^{-1}H_{YX}\tilde{Y}'\right] \\ &= P_XP_Y.\end{aligned}$$

If  $p = 1$ , the decomposition  $Q = S\Lambda T'$  is a singular value decomposition of  $Q$ . In this case we know that  $Q$  has one singular value which equals 0,  $k - 2$  singular values which equal 1, and a singular value which immediately follows from (6.1):

$$\lambda_2 = \|x\| \|y\|.$$

Since  $x'y = 1$ , obviously,  $\|x\| \|y\| \geq 1$  and the equality sign holds if and only if  $Q$  is symmetric (so  $Q = P_X$ ). To summarize, we have the following result.



**Corollary 1.** *Let  $x, y$  be  $k \times 1$  vectors, such that  $y'x = 1$  and  $\text{rank}(x, y) = 2$ . If we define  $r_{xy} = (x'y)/(\|x\| \|y\|) = \|x\|^{-1} \|y\|^{-1}$ , then a singular value decomposition of  $Q$  is  $S\Lambda T'$ , where*

$$\begin{aligned}\Lambda &= \text{diag}(0, \|x\| \|y\|, 1, \dots, 1) \\ S &= (\|y\|^{-1}y, \|x\|^{-1}(1 - r_{xy}^2)^{-\frac{1}{2}}(x - \|y\|^{-2}y), w_1, \dots, w_{k-2}) \\ T &= (\|x\|^{-1}x, -\|y\|^{-1}(1 - r_{xy}^2)^{-\frac{1}{2}}(y - \|x\|^{-2}x), w_1, \dots, w_{k-2}),\end{aligned}$$

*such that  $S$  and  $T$  are orthogonal matrices. The vectors  $w_1, \dots, w_{k-2}$  have unit length, are mutually orthogonal and are also orthogonal to  $x$  and  $y$ .*

Observe Gram-Schmidt orthogonalization in this corollary: in  $L_{x,y} = \{\alpha x + \beta y | \alpha, \beta \in \mathbb{R}\}$  find the vector orthogonal to  $y$ , resp.  $x$ .

## 7 A semi-eigenvalue decomposition

In this section, we will derive a decomposition for the matrix  $Q$  which is somewhat similar to an eigenvalue decomposition. As opposed to the previous section, the equality  $Y'X = I_p$  we assumed throughout need not hold. It turns out that, under particular conditions, we can write

$$Q = UDU^{-1}, \quad (7.1)$$

where  $D$  is a block-diagonal matrix. Equation (7.1) shows that  $Q$  is similar to a block-diagonal matrix. Although the matrix  $D$  has a simple structure, it does not give us the eigenvalues of  $Q$ , like the spectral decomposition does. Moreover, the matrix  $U$  need not be orthogonal. We will show, however, that in some specific situations equation (7.1) gives an eigenvalue decomposition of  $Q$ . We will only discuss the case where  $X$  and  $Y$  span different spaces. Section 7.1 deals with the general case, whereas section 7.2 focuses on the case  $p = 1$ .

### 7.1 $X$ and $Y$ span different spaces

Consider the matrix  $Q = I_k - XY'$ , where  $X$  and  $Y$  are  $k \times p$  matrices of full column rank,  $R(X) \cap R(Y) = \{0\}$ . Note once more that we do not restrict ourselves to the case that  $Y'X = I_p$ . The aim is to find a decomposition  $QU = UD$ , where  $U$  is nonsingular and  $D$  has a simple structure. Because  $QX = X(I_p - Y'X)$ , the matrix  $X$  is a natural candidate to be part of  $U$ . We prefer to normalize the columns of  $U$ , which, in the

notation of section 6, leads to

$$\begin{aligned} Q\tilde{X} &= \tilde{X}(I_p - (X'X)^{\frac{1}{2}}Y'\tilde{X}) \\ &= \tilde{X}(I_p - (X'X)^{\frac{1}{2}}(Y'Y)^{\frac{1}{2}}\tilde{Y}'\tilde{X}). \end{aligned} \quad (7.2)$$

On the other hand,  $Qa = a$  for all  $a \in \mathbb{R}^k$  with  $Y'a = 0$ , and we see that all vectors orthogonal to the columns of  $Y$  are eigenvectors of  $Q$  with eigenvalue 1.

Now we can construct the decomposition  $QU = UD$  as follows. Let  $U = (U_1, U_2, U_3)$  and  $D = \text{diag}(D_1, D_2, D_3)$ , then (7.2) suggests to take  $U_1 = \tilde{X}$  and  $D_1 = I_p - (X'X)^{\frac{1}{2}}(Y'Y)^{\frac{1}{2}}\tilde{Y}'\tilde{X}$ . If we consider the space spanned by the columns of  $\tilde{X}$  and  $\tilde{Y}$ , we are now looking for vectors orthogonal to the columns of  $Y$ , which leads us to the choice  $U_2 = P_Y\tilde{X}(\tilde{X}'P_Y\tilde{X})^{-\frac{1}{2}} = P_Y\tilde{X}H_{XY}^{-\frac{1}{2}}$ , because then

$$QU_2 = QP_Y\tilde{X}H_{XY}^{-\frac{1}{2}} = P_Y\tilde{X}H_{XY}^{-\frac{1}{2}} = U_2, \quad (7.3)$$

and  $U_2'U_2 = I_p$ . The columns of  $U_2$  are eigenvectors of  $Q$  with eigenvalue 1, so that  $D_2 = I_p$ . It is not immediately apparent that the columns of  $U_1$  and  $U_2$  span different spaces, that is  $\text{rank}(U_1, U_2) = 2p$ , which is a necessary condition for  $U$  to be nonsingular. Because  $\text{rank}(U_1, U_2) = \text{rank}(U_1, U_2)'(U_1, U_2)$ , we can also consider the matrix

$$(U_1, U_2)'(U_1, U_2) = \begin{pmatrix} I_p & H_{XY}^{\frac{1}{2}} \\ H_{XY}^{\frac{1}{2}} & I_p \end{pmatrix}. \quad (7.4)$$

From (7.4) we see that

$$\begin{aligned} |(U_1, U_2)'(U_1, U_2)| &= |I_p - H_{XY}| = |\tilde{X}'\tilde{Y}\tilde{Y}'\tilde{X}| \\ &= |(X'X)^{-\frac{1}{2}}X'Y(Y'Y)^{-1}Y'X(X'X)^{-\frac{1}{2}}|, \end{aligned}$$

so that  $|(U_1, U_2)'(U_1, U_2)| = 0$  if and only if  $|X'Y| \neq 0$ , which means that  $X'Y$  must be nonsingular. Thus, if  $X'Y$  is nonsingular,  $(U_1, U_2)'(U_1, U_2)$  is of full rank, and therefore, so is  $(U_1, U_2)$ . We see that in this case the columns of  $U_1$  and  $U_2$  constitute an independent system of  $2p$  vectors in  $\mathbb{R}^k$ . Now we can find vectors  $c_1, \dots, c_{k-2p}$  that are mutually orthogonal, have unit length, and are perpendicular to the columns of  $X$  and  $Y$  and therefore also orthogonal to the columns of  $U_1$  and  $U_2$ . If we define  $C = (c_1, \dots, c_{k-2p})$ , then  $QC = C$ , which means that the columns of  $C$  are eigenvectors of  $Q$  with eigenvalue 1. With  $U_3 = C$  and  $D_3 = I_{k-2p}$ , our matrix decomposition  $QU = UD$  is completed. Note that the matrix  $V$  consists of vectors of unit length which are mutually orthogonal, except for the vectors in  $U_1$  and  $U_2$ .

To determine  $U^{-1}$ , note from (7.1) that

$$Q'(U^{-1})' = (U^{-1})'D',$$

so that finding a similarity representation for  $Q'$  leads to interchanging the role of  $X$  and  $Y$ . This leads to the guess

$$(U^{-1})' = (\tilde{Y}\Lambda, P_X\tilde{Y}\Gamma, C),$$

where  $\Lambda$  and  $\Gamma$  are  $p \times p$  matrices. Now  $V^{-1}V = I_k$  implies

$$\begin{pmatrix} \Lambda'\tilde{Y}' \\ \Gamma'\tilde{Y}'P_X \\ C' \end{pmatrix} \begin{pmatrix} \tilde{X} & P_Y\tilde{X}H_{XY}^{-\frac{1}{2}} & C \end{pmatrix} = I_k.$$

The off-diagonal blocks are indeed equal to zero, whereas  $\Lambda$  and  $\Gamma$  must satisfy the equations

$$\Lambda'\tilde{Y}'\tilde{X} = I_p, \quad (7.5)$$

and

$$\begin{aligned} \Gamma'\tilde{Y}'P_X P_Y\tilde{X}H_{XY}^{-\frac{1}{2}} &= \Gamma'\tilde{Y}'(I_k - \tilde{X}\tilde{X}' - \tilde{Y}\tilde{Y}' + \tilde{X}\tilde{X}'\tilde{Y}\tilde{Y}')\tilde{X}H_{XY}^{-\frac{1}{2}} \\ &= -\Gamma'(\tilde{Y}'\tilde{X} - \tilde{Y}'\tilde{X}\tilde{X}'\tilde{Y}\tilde{Y}'\tilde{X})H_{XY}^{-\frac{1}{2}} \\ &= -\Gamma'\tilde{Y}'\tilde{X}H_{XY}^{\frac{1}{2}}. \end{aligned} \quad (7.6)$$

From (7.5), respectively (7.6), we see that

$$\begin{aligned} \Lambda &= (\tilde{X}'\tilde{Y})^{-1} \\ \Gamma &= -(\tilde{X}'\tilde{Y})^{-1}H_{XY}^{-\frac{1}{2}}. \end{aligned}$$

Therefore,

$$(U^{-1})' = \left( \tilde{Y}(\tilde{X}'\tilde{Y})^{-1}, -P_X\tilde{Y}(\tilde{X}'\tilde{Y})^{-\frac{1}{2}}H_{XY}^{-\frac{1}{2}}, C \right). \quad (7.7)$$

We summarize the results in the following theorem.

**Theorem 5.** *Let  $X, Y$  be  $k \times p$  matrices of rank  $p$ , such that  $R(X) \cap R(Y) = \{0\}$ , and  $X'Y$  is nonsingular. Define  $\tilde{X} = X(X'X)^{-\frac{1}{2}}$ ,  $\tilde{Y} = Y(Y'Y)^{-\frac{1}{2}}$ ,  $H_{XY} = \tilde{X}P_Y\tilde{X}$  and  $H_{YX} = \tilde{Y}P_X\tilde{Y}$ . Then  $Q = I_k - XY'$  can be written as  $Q = UDU^{-1}$ , where*

$$\begin{aligned} D &= \text{diag}(I_p - (X'X)^{\frac{1}{2}}(Y'Y)^{\frac{1}{2}}\tilde{Y}'\tilde{X}, I_p, I_{k-2p}) \\ U &= (\tilde{X}, P_Y\tilde{X}H_{XY}^{-\frac{1}{2}}, C) \\ (U^{-1})' &= \left( \tilde{Y}(\tilde{X}'\tilde{Y})^{-1}, -P_X\tilde{Y}(\tilde{X}'\tilde{Y})^{-\frac{1}{2}}H_{XY}^{-\frac{1}{2}}, C \right), \end{aligned}$$

such that the  $k \times (k - 2p)$  matrix  $C$  has the property that  $C'C = I_{k-2p}$  and  $Y'C = X'C = 0$ .

If  $Y'X = I_p$ , theorem 5 gives us an eigenvalue decomposition of  $Q$ :

**Corollary 2.** Let  $X, Y$  be  $k \times p$  matrices of rank  $p$ , such that  $Y'X = I_p$ ,  $R(X) \cap R(Y) = \{0\}$ . Define  $\tilde{X} = X(X'X)^{-\frac{1}{2}}$ ,  $\tilde{Y} = Y(Y'Y)^{-\frac{1}{2}}$ ,  $H_{XY} = \tilde{X}P_Y\tilde{X}$  and  $H_{YX} = \tilde{Y}P_X\tilde{Y}$ . Then  $QU = UD$  is an eigenvalue decomposition of  $Q = I_k - XY'$ , where

$$\begin{aligned} D &= \text{diag}(0_p, I_{k-p}) \\ U &= (\tilde{X}, P_Y\tilde{X}H_{XY}^{-\frac{1}{2}}, C) \\ (U^{-1})' &= \left( Y(X'X)^{\frac{1}{2}}, -P_XY(X'X)^{\frac{1}{2}}H_{XY}^{-\frac{1}{2}}, C \right), \end{aligned}$$

such that the  $k \times (k - 2p)$  matrix  $C$  has the property that  $C'C = I_{k-2p}$  and  $Y'C = X'C = 0$ .

Corollary 2 shows, in correspondence with (1.8), that  $Q$  has  $p$  eigenvalues equal to 0, and  $k - p$  eigenvalues equal to 1. Moreover, corollary 2 also gives us the corresponding eigenvectors.

## 7.2 The case $p = 1$

Consider the matrix  $Q = I_k - xy'$ , where  $x$  and  $y$  are  $k \times 1$  vectors. Some interesting observations are:

- (a)  $Qx = (1 - y'x)x$ ;
- (b)  $Qa = a$ , for all  $a$  with  $y'a = 0$ ;
- (c)  $Qa = 0$  for some  $a \neq 0$  if and only if  $x'y = 1$ ;
- (d)  $\text{rank}(Q) = k$  for  $x'y \neq 1$  and  $\text{rank}(Q) = k - 1$  if  $x'y = 1$ ;
- (e)  $Q$  is symmetric if and only if  $y = \lambda x$  for some  $\lambda \neq 0$  (and hence  $x'y \neq 0$ );
- (f)  $Q^2 = I_k + (x'y - 2)xy'$ , hence  $Q$  is idempotent if and only if  $x'y = 1$ .

*Proof of (c) and (e).*

(c)  $Qa = 0$  for some  $a \neq 0$  if and only if  $Q$  is singular. Because  $|Q| = |I_k - xy'| = 1 - y'x$ , this holds if and only if  $x'y = 1$ .

(e) If  $y = \lambda x$ , then obviously  $Q$  is symmetric. Conversely, if  $Q$  is symmetric, then  $Qx = Q'x$  so that  $(y'x)x = (x'x)y$  where  $y'x \neq 0$ , and  $y = (y'x)(x'x)^{-1}x$ .

With regard to the eigenvalues of  $Q$  we now have to distinguish two cases:

- (i)  $x'y \neq 0$ : From (a) it is obvious that  $Q$  has an eigenvector  $x$  with eigenvalue  $1 - x'y \neq 1$  with multiplicity 1. From (b) we see  $Q$  has an eigenvalue 1 with multiplicity  $k - 1$ . The eigenspace corresponding to this eigenvalue is  $L_y^\perp = \{a \in \mathbb{R}^k | a'y = 0\}$ . Because  $x$  is not perpendicular to  $y$ , there are  $k$  independent eigenvectors. So,  $Q$  is similar to a diagonal matrix.
- (ii)  $x'y = 0$ : From (1.7) we know that  $Q$  has one eigenvalue equal to 1 with multiplicity  $k$ . We know from (b) that all vectors  $c$  with  $c'y = 0$  are eigenvectors of  $Q$  with eigenvalue 1. Suppose that  $\alpha y + \beta c$  is an eigenvector with  $c'y = 0$  and  $\alpha, \beta \in \mathbb{R}$ . Then  $\alpha y + \beta c = Q(\alpha y + \beta c) = \alpha y + \beta c - \alpha y y' x$  so that  $\alpha = 0$ . Therefore, all eigenvectors of  $Q$  are orthogonal to  $y$  and the eigenspace of  $Q$  is  $L_y^\perp$  and has dimension  $k - 1$ . In this case,  $Q$  is not similar to a diagonal matrix.

Just like the case  $p = 1$  in section 6 gave us the singular values of  $Q$ , the case  $p = 1$  now gives us the eigenvalues.

**Corollary 3.** *Let  $x, y$  be  $k \times 1$  vectors, such that  $x'y \neq 0$  and  $\text{rank}(x, y) = 2$ . If we define  $r_{xy} = (x'y)/(\|x\| \|y\|) = \|x\|^{-1} \|y\|^{-1}$ , then  $Q = U D U^{-1}$ , where  $D$  is a diagonal matrix with the eigenvalues of  $Q$  along its diagonal*

$$\begin{aligned} D &= \text{diag}(1 - x'y, 1, 1, \dots, 1) \\ U &= (\|x\|^{-1}x, \|x\|^{-1}(1 - r_{xy}^2)^{-\frac{1}{2}}(x - \frac{x'y}{y'y}y), c_1, \dots, c_{k-2}), \\ (U^{-1})' &= \left( r_{xy}^{-1} \|y\|^{-1} y, -r_{xy}^{-1} (1 - r_{xy}^2)^{-\frac{1}{2}} \|y\|^{-1} (y - \frac{x'y}{x'x}x), c_1, \dots, c_{k-2} \right). \end{aligned}$$

where  $c_1, \dots, c_{k-2}$  are mutually orthogonal, have unit length and they are perpendicular to  $x$  and  $y$ .

The second vector of  $U$  is chosen such that it is in the space  $L_{x,y} = \{\alpha x + \beta y | \alpha, \beta \in \mathbb{R}\}$  and perpendicular to  $y$ . Here we used Gram-Schmidt orthogonalization. Note once more that the matrix  $U$  consists of vectors of unit length which are mutually orthogonal except for the first and the second vector. If these vectors were also orthogonal, we would have  $U^{-1} = U'$  and hence  $Q = Q'$ , which is not the case. However, if  $Q$  is symmetric, then the vectors  $x$  and  $y$  span the same space, so that  $U$  is orthogonal and corollary 3 gives the spectral decomposition of  $Q$ .

## 8 Eigenvalues and eigenvectors of $D - xy'$

In this section, we will study the eigenvalues and eigenvectors of the matrix  $D - xy'$ , where  $D = \text{diag}(d)$  is a nonsingular diagonal matrix and  $x$  and  $y$  are  $k \times 1$  vectors. Similar problems have been studied by Vermeulen (1967), Klamkin (1970), Wansbeek (1985), Watson (1996) and Trenkler (2000). Trenkler notes that we need not restrict ourselves to diagonal  $D$ . If we look at the eigenvalues of  $A - xy'$ , where  $A$  is a nonsingular symmetric matrix, then there exists an orthogonal matrix  $H$ , such that  $A = H\Lambda H'$ , where  $\Lambda$  is a nonsingular diagonal matrix with the eigenvalues of  $A$  along its diagonal. Since  $A - xy'$  and  $\Lambda - H'xy'H$  have the same eigenvalues, we might as well study the matrix  $\Lambda - xy'$ . The condition of symmetry can be replaced by the requirement that  $A$  is similar to a diagonal matrix  $\Lambda$ , that is,  $A = U\Lambda U^{-1}$  for some  $k \times k$  matrix  $U$ . In this case, the eigenvalues of  $A - xy'$  coincide with those of  $\Lambda - U^{-1}xy'U$ .

### 8.1 Eigenvalues

Consider the matrix  $D - xy'$ , where  $D = \text{diag}(d)$  is a nonsingular diagonal matrix and  $d$ ,  $x$  and  $y$  are  $k \times 1$  vectors. We are interested in the eigenvalues of this matrix. Inspired by Vermeulen (1967), we now present the following theorem.

**Theorem 6.** *If  $D = \text{diag}(d)$  is a nonsingular diagonal  $k \times k$  matrix and  $x$  and  $y$  are  $k \times 1$  vectors with  $x_i y_i \neq 0, i = 1, \dots, k$ , then*

$$|D - xy'| = (-1)^{s(x,y)} |D_{xy} - vv'|,$$

where

$$\begin{aligned} S_x &= \text{diag}(\text{sgn } x_1, \dots, \text{sgn } x_k) \\ S_y &= \text{diag}(\text{sgn } y_1, \dots, \text{sgn } y_k) \\ D_{xy} &= DS_x S_y \\ v &= (|x_1 y_1|^{\frac{1}{2}}, \dots, |x_k y_k|^{\frac{1}{2}})' \\ s(x, y) &= \#\{i = 1, \dots, k \mid x_i y_i < 0\}. \end{aligned}$$

*Proof.* Let  $D_x = \text{diag}(|x_1|^{\frac{1}{2}}, \dots, |x_k|^{\frac{1}{2}})$  and  $D_y = \text{diag}(|y_1|^{\frac{1}{2}}, \dots, |y_k|^{\frac{1}{2}})$ . Then

$$\begin{aligned}
|D - xy'| &= |D_x| |D_x^{-1} D D_y^{-1} - D_x^{-1} xy' D_y^{-1}| |D_y| \\
&= |D_y D_x^{-1} D D_y^{-1} D_x - D_y D_x^{-1} xy' D_y^{-1} D_x| \\
&= |D - S_x v v' S_y| \\
&= |S_x| |S_x D S_y - v v'| |S_y| \\
&= (-1)^{s(x,y)} |D_{xy} - v v'|.
\end{aligned}$$

□

If, for some index  $j$ , we have  $x_j y_j = 0$ , then we can expand  $|D - xy'|$  along its  $j^{\text{th}}$  row or column:

$$|D - xy'| = d_j |D_{jj} - (xy')_{jj}|,$$

where  $D_{jj}$  and  $(xy')_{jj}$  denote the matrices obtained by deleting the  $j^{\text{th}}$  row and the  $j^{\text{th}}$  column of  $D$ , respectively  $xy'$ . We can continue this process until none of the  $x_i y_i = 0$  and then apply theorem 6 to the remaining part of the matrix  $D - xy'$ .

From theorem 6, we see that  $|D - xy'|$  is the same as  $|D_{xy} - v v'|$ , except possibly for a difference in sign. Likewise,

$$|\lambda I_k - (D - xy')| = (-1)^{s(x,y)} |(\lambda I_k - D) S_x S_y + v v'| \quad (8.1)$$

with  $S_x, S_y$  and  $v$  as defined in theorem 6. Equation (8.1) implies that if  $S_x S_y = I_k$ , that is if  $x_i y_i > 0$  for  $i = 1, \dots, k$ , then the roots of the characteristic equation of  $D - xy'$  and those of the symmetric matrix  $D - v v'$  are the same. If, on the other hand  $S_x S_y = -I_k$ , that is if  $x_i y_i < 0$  for  $i = 1, \dots, k$ , then the roots of the characteristic equation  $D - xy'$  and those of the symmetric matrix  $D + v v'$  are the same. Because the roots of a symmetric matrix are always real, we have shown that if all  $x_i y_i$  have the same sign, then the eigenvalues of  $D - xy'$  are real.

**Theorem 7.** *If  $D = \text{diag}(d)$  is a nonsingular diagonal  $k \times k$  matrix and  $x, y$  are  $k \times 1$  vectors such that  $x_i y_i < 0$  for  $i = 1, \dots, k$  or  $x_i y_i > 0$  for  $i = 1, \dots, k$ , then  $Q = D - xy'$  has real eigenvalues.*

Vermeulen (1967) showed that the roots of the determinantal equation

$$|\lambda I_k + \text{diag}(e) + a_i' l_i| \quad (8.2)$$

with  $a_i$  and  $e_i$  strictly positive, are real by using a similar argument as used in theorem 6. By constructing a difference equation for the determinantal equation, he also showed that these roots are negative. These results immediately follow from theorem 6, because

$$|\lambda I_k + \text{diag}(e) + a_i' l_i| = |\lambda I_k + \text{diag}(e) + v v'|,$$

with  $v = (\sqrt{a_1}, \dots, \sqrt{a_k})'$ , so that the eigenvalues of  $\text{diag}(e) + a_i'_{i,k}$  are the same as the eigenvalues of the symmetric matrix  $\text{diag}(e) + vv'$ . These eigenvalues are positive, because  $\text{diag}(e) + vv'$  is positive definite. This implies that the roots of the determinantal equation (8.2) are real and negative.

Trenkler (2000) remarks in his paper that the matrix  $T = A + ab'$ , with  $A$  being symmetric and nonsingular, has always real eigenvalues. However, according to the condition of theorem 7 that all  $a_i b_i$  should have the same sign is vital more or less, as can be seen from the following example.

**Example 1.** Consider the matrix  $D - xy'$  with  $D = \text{diag}(1, 2)$ ,  $x = (-1, -1)'$  and  $y = (1, -3)'$ . It is easily derived that in this case  $|\lambda I_2 - (D - xy')| = \lambda^2 - \lambda + 1$ , so that both eigenvalues are complex.  $\square$

The eigenvalues of  $D - xy'$  can be determined from the characteristic equation

$$\begin{aligned}
0 &= |\lambda I_k - (D - xy')| \\
&= |(\lambda I_k - D)(I_k + (\lambda I_k - D)^{-1}xy')| \\
&= |\lambda I_k - D|(1 + y'(\lambda I_k - D)^{-1}x) \\
&= \left[ \prod_{i=1}^k (\lambda - d_i) \right] \left[ 1 + \sum_{i=1}^k \frac{x_i y_i}{\lambda - d_i} \right] \\
&= \prod_{i=1}^k (\lambda - d_i) + \sum_{i=1}^k x_i y_i \prod_{j \neq i} (\lambda - d_j) \tag{8.3}
\end{aligned}$$

The following theorem immediately follows from (8.3). It covers the theorems 2, 3, and 4 of Trenkler (2000).

**Theorem 8.** Consider  $D - xy'$ , where  $D = \text{diag}(d)$  is a nonsingular diagonal matrix and  $x$  and  $y$  are  $k \times 1$  vectors.

(i) If all  $d_i$  are different and all  $x_i y_i \neq 0$ , then none of the  $d_i$  is an eigenvalue of  $D - xy'$ . In this case,  $\lambda$  is an eigenvalue if and only if

$$1 + \sum_{i=1}^k \frac{x_i y_i}{\lambda - d_i} = 0. \tag{8.4}$$

(ii) If all  $d_i$  are different, but, for index  $j$ , we have  $x_j y_j = 0$ , then  $d_j$  is an eigenvalue of  $D - xy'$ .



(iii) If some of the  $d_i$ 's coincide,  $d_i$  is an eigenvalue of  $D - xy'$ .

Note that we can find all eigenvalues of  $D - xy'$  by combining the different cases considered in this theorem. In the most general case where some of the  $d_i$  coincide, some of the  $x_i y_i$  are equal to zero but indices  $i$  also exist such that  $d_i$  has multiplicity 1 and  $x_i y_i \neq 0$ , the procedure is as follows. Partition  $D$  in blocks of ascending size:

$$D = \text{diag}(d_1 I_{k_1}, d_2 I_{k_2}, \dots, d_k I_{k_r}) \quad (8.5)$$

where  $\sum_{j=1}^r k_j = k$ . Partition  $x$  and  $y$  accordingly:

$$\begin{aligned} x &= (x_{k_1}, x_{k_2}, \dots, x_{k_r})' \\ y &= (y_{k_1}, y_{k_2}, \dots, y_{k_r})'. \end{aligned} \quad (8.6)$$

First, suppose  $d_j$  has multiplicity  $k_j > 1$ . Then we know from theorem 8 that  $d_j$  is an eigenvalue of  $D - xy'$ . Equation (8.3) then gives

$$\begin{aligned} |\lambda I_k - (D - xy')| &= \prod_{j=1}^r (\lambda - d_j)^{k_j} + \sum_{j=1}^r (x'_{k_j} y_{k_j}) (\lambda - d_j)^{k_j-1} \prod_{h \neq j} (\lambda - d_h)^{k_h} \\ &= \left[ \prod_{j=1}^r (\lambda - d_j)^{k_j-1} \right] \left[ \prod_{j=1}^r (\lambda - d_j) + \right. \\ &\quad \left. + \sum_{j=1}^r (x'_{k_j} y_{k_j}) \prod_{h \neq j} (\lambda - d_h) \right] \end{aligned} \quad (8.7)$$

From (8.7) we observe that in this case  $\lambda = d_j$  has multiplicity  $k_j - 1$ . Moreover, the remaining eigenvalues can be found by putting the second factor on the right-hand side of (8.7) equal to zero. The equation to be solved is then exactly of the type (8.3), so that the remaining eigenvalues can be determined from (8.4). Note that  $\lambda = d_j$  can have multiplicity  $k_j$  if and only if  $x'_{k_j} y_{k_j} = 0$ .

Second, consider the set of indices  $J$  for which the  $x_j y_j$  equal zero and  $d_j$  has multiplicity 1, that is,  $J = \{j : x_j y_j = 0, d_j \text{ has multiplicity } 1\}$ . Then we can derive

$$|\lambda I_k - (D - xy')| = \prod_{j \in J} (\lambda - d_j) |\lambda I_{k^*} - (\tilde{D} - \tilde{x} \tilde{y}')|, \quad (8.8)$$

where  $\tilde{D}$  is the matrix obtained from  $D$  by deleting the  $j^{\text{th}}$  row and column for all  $j \in J$ ,  $\tilde{x}$  and  $\tilde{y}$  are the vectors obtained from  $x$ , respectively  $y$ , by deleting the  $j^{\text{th}}$  element for all  $j \in J$ , and  $k^* = |J|$ , the number of elements in  $J$ . Equation (8.8) shows that all  $d_j$

with  $j \in J$  are unique eigenvalues of  $D - xy'$ . Moreover, to determine the remaining eigenvalues of  $D - xy'$ , we can apply part (i) of theorem 8 to the matrix  $\tilde{D} - \tilde{x}\tilde{y}'$  and we are finished.

If all  $d_i$  coincide, that is, if  $D = dI_k$ , then equation (8.3) gives

$$|\lambda I_k - (D - xy')| = (\lambda - d)^{k-1} \left[ (\lambda - d) + \sum_{i=1}^k x_i y_i \right] = 0,$$

so that  $\lambda = d$  is an eigenvalue with multiplicity  $n - 1$  and the remaining eigenvalue equals  $d - xy'$ .

In the situation that all  $d_i$  are different and all  $x_i y_i \neq 0$ , case (i) of theorem 8, the eigenvalues of  $D - xy'$  must be determined by solving equation (8.4). Note that, in this case,  $\lambda = 0$  is a solution of (8.4) if and only if  $y'D^{-1}x = 1$ , that is,  $D - xy'$  is singular. Moreover, if  $y'D^{-1}x = 1$ , equation (8.3) simplifies to

$$\begin{aligned} 0 &= |\lambda I_k - (D - xy')| \\ &= |\lambda I_k - D| (y'(\lambda I_k - D)D^{-1}(\lambda I_k - D)^{-1}x + y'(\lambda I_k - D)^{-1}x) \\ &= |\lambda I_k - D| (y'(\lambda D^{-1} - I_k)(\lambda I_k - D)^{-1}x + y'(\lambda I_k - D)^{-1}x) \\ &= |\lambda I_k - D| \lambda y'D^{-1}(\lambda I_k - D)^{-1}x. \end{aligned}$$

Thus, apart from equation (8.4), the eigenvalues different from zero also satisfy

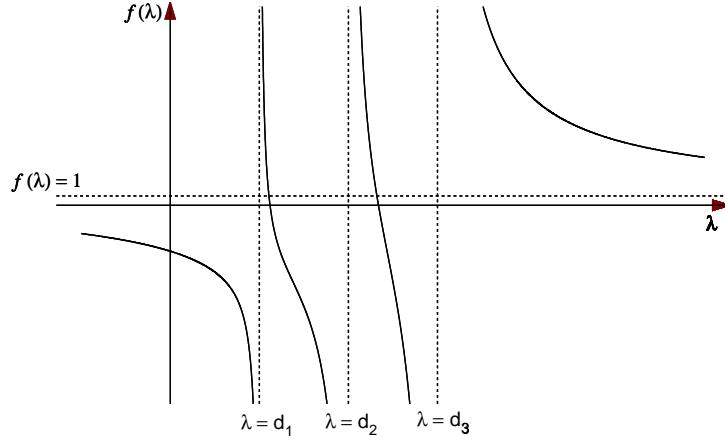
$$\sum_{i=1}^k \frac{x_i y_i}{d_i(\lambda - d_i)} = 0.$$

From now on we assume, without loss of generality, that the  $d_i$  are arranged in ascending order. Just as Klamkin (1970), Wansbeek (1985) and Trenkler (2000), we want to pay attention to the location of the eigenvalues for this special case. If we want to say something about the location of the eigenvalues, we would like them to be real, and therefore we restrict ourselves to the situation in which all  $x_i y_i$  have the same sign. In line with Klamkin (1970), consider the graph of

$$f(\lambda) = 1 + \frac{x_1 y_1}{\lambda - d_1} + \frac{x_2 y_2}{\lambda - d_2} + \dots + \frac{x_n y_n}{\lambda - d_k}. \quad (8.9)$$

This graph is continuous except at the points  $\lambda = d_1, d_2, \dots, d_k$ , which correspond to vertical asymptotes. It follows by continuity that there are  $k$  real roots such that between every two successive  $d_i$  lies an eigenvalue, that is

$$d_1 < \lambda_1 < d_2 < \lambda_2 < \dots < \lambda_{k-1} < d_k < \lambda_k. \quad (8.10)$$



**Figure 1.** Example of the roots of the characteristic equation  $|\lambda I_k - (D - xy')| = 0$  for  $k = 3$  and  $x_i y_i > 0$ ,  $i = 1, 2, 3$ .

A typical graph for  $k = 3$  is shown in figure 1.

In case of situation (ii) or (iii) of theorem 8, where some of the  $d_j$  coincide or  $x_j y_j = 0$ , we know that  $d_j$  is an eigenvalue of  $D - xy'$ . The other eigenvalues are located as before, so that in both these situations

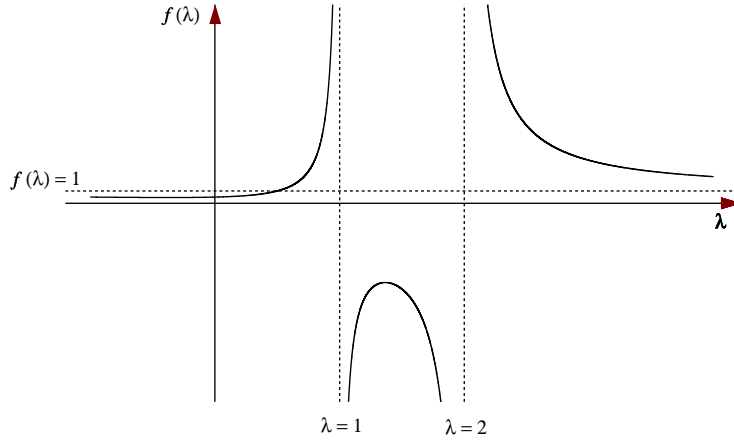
$$d_1 \leq \lambda_1 \leq d_2 \leq \lambda_2 \leq \dots \leq \lambda_{k-1} \leq d_k \leq \lambda_k. \quad (8.11)$$

To show that we cannot locate the eigenvalues among the  $d_i$  as easily as in equation (8.11) if we do not impose any restrictions on the sign of the  $x_i y_i$ , consider the following example.

**Example 2.** We continue with example 1, and construct the graph of (8.9), that is, the graph of

$$f(\lambda) = 1 - \frac{1}{\lambda - 1} + \frac{3}{\lambda - 2}.$$

We know that there are no real roots, so that the graph never intersects the  $x$ -axis. The corresponding graph is shown in figure 2.  $\square$



**Figure 2.** Graph corresponding to example 2.

## 8.2 Eigenvectors

Watson (1996) considered the eigenvectors of the matrix  $P - pp'$ , with  $P = \text{diag}(p)$  and  $\sum p_i = 1$ . Somewhat more general, we will consider in this section the eigenvectors of  $D - xy'$ .

We begin with the case where all  $d_i$  are different and all  $x_i y_i \neq 0$ , case (i) of theorem 8. The elements of the eigenvector  $v_j = (v_{1j}, \dots, v_{kj})'$  of  $D - xy'$  corresponding to  $\lambda_j$  must satisfy

$$d_i v_{ij} - (y' v_j) x_i = \lambda_j v_{ij}, \quad i = 1, \dots, k,$$

so that

$$v_{ij} = (y' v_j) \frac{x_i}{d_i - \lambda_j}. \quad (8.12)$$

Note that  $\lambda_j \neq d_i$ . In order to have a true eigenvector, we must have  $y' v_j \neq 0$ . If we choose

$$v_{ij} = \frac{x_i}{d_i - \lambda_j} \neq 0, \quad (8.13)$$

then

$$y' v_j = \sum_{i=1}^k \frac{x_i y_i}{d_i - \lambda_j} = 1,$$

because  $\lambda_j$  satisfies (8.4). This shows that we can indeed find the elements of the eigenvector  $v_j$  by means of equation (8.13). If we assume additionally that the  $d_i$  are ordered and that all  $x_i y_i$  have the same sign, then equality (8.10) holds, so that the  $v_{1j}, \dots, v_{jj}$  have the same sign which is opposite to the sign of the  $v_{j+1,j}, \dots, v_{kj}$ . Note that we do not need the vector  $y$  to determine the eigenvectors of  $D - xy'$ , this vector is only relevant for determining the eigenvalues of  $D - xy'$ .

If all  $d_i$  are different, but one or more of the  $x_j y_j = 0$ , the second case of theorem 8, then  $d_j$  is an eigenvalue of  $D - xy'$ . First, consider the situation that  $y_j = 0$  and  $x_j$  is arbitrary. By using a similar approach as above, it is straightforward to show that in this case  $e_j$ , the  $j^{\text{th}}$  unit vector, is an eigenvector corresponding to  $d_j$ . If, on the other hand  $x_j = 0$  and  $y_j \neq 0$ , then the elements of the eigenvector  $v = (v_1, \dots, v_k)'$  corresponding to  $d_j$  must satisfy

$$d_i v_i - (y'v)x_i = d_j v_i, \quad i = 1, \dots, k. \quad (8.14)$$

For  $i \neq j$  this leads to

$$v_i = \frac{(y'v)x_i}{d_i - d_j},$$

so that

$$y'v = y_j v_j + (y'v) \sum_{i \neq j} \frac{x_i y_i}{d_i - d_j},$$

and  $v_j$  follows:

$$v_j = \frac{y'v}{y_j} \left(1 - \sum_{i \neq j} \frac{x_i y_i}{d_i - d_j}\right).$$

Because  $y'v$  is a constant factor, this shows that we can choose  $v$  such that

$$\begin{aligned} v_i &= \frac{x_i}{d_i - d_j} \quad \text{for } i \neq j \\ v_j &= \frac{1}{y_j} \left(1 - \sum_{i \neq j} \frac{x_i y_i}{d_i - d_j}\right). \end{aligned}$$

Note that  $y'v = 1$  with this choice of  $v$ .

For case (iii) of theorem 8, where some of the  $d_i$  coincide, partition  $D$ ,  $x$  and  $y$  as in (8.5) and (8.6). We know from theorem 8 that if  $k_j > 1$ , then  $d_j$  is an eigenvalue of  $D - xy'$ . We partition an eigenvector  $v$  corresponding to  $d_j$  in a similar fashion as in (8.6). Assume that  $x_{k_j}, y_{k_j} \neq 0$ . It is easy to show that in this case,  $v = (v_{k_1}, v_{k_2}, \dots, v_{k_r})'$  consists of zeros, except for the subvector  $v_{k_j}$ . This subvector must satisfy  $y'_{k_j} v_{k_j} = 0$ . This implies that in this case,  $d_j$  has  $k_j - 1$  eigenvectors  $(0, \dots, 0, v_{k_j}, 0, \dots, 0)'$ , where the  $v_{k_j}$  are orthogonal to  $y_{k_j}$  and are also mutually orthogonal.

## References

- Bouillion, T. L. and P. L. Odell (1971), *Generalized Inverse Matrices*, John Wiley & Sons, Inc., New York.
- De Boer, P. M. C. and R. Harkema (1984), “Maximum likelihood estimation of sum-constrained linear models when samples are small”, *Kwantitatieve Methoden*, **16**, 97–108.
- Dol, W. (1991), *Small Area Estimation; a Synthesis Between Sampling Theory and Econometrics*, PhD thesis, University of Groningen, Faculty of Economics, Groningen.
- Dol, W., T. Steerneman, and T. Wansbeek (1996), “Matrix algebra and sampling theory: The case of the Horvitz-Thompson estimator”, *Linear Algebra and its Applications*, 225–238.
- Klamkin, M. S. (1970), “On the roots of a certain determinantal equation”, *Mathematical Gazette*, **54**(387), 57–58.
- Marsaglia, G. and G. P. H. Styan (1974), “Rank conditions for generalized inverses of partitioned matrices”, *Sankhyā, Series A*, **36**, 347–442.
- Neudecker, H. (1995), “Mathematical properties of the variance of the multinomial distribution”, *Journal of Mathematical Analysis and Applications*, **189**, 757–762.
- Ouellette, D. V. (1981), “Schur complements and statistics”, *Linear Algebra and its Applications*, **36**, 187–295.
- Rao, C. R. and S. J. Mitra (1971), *Generalized Inverse of Matrices and its Applications*, John Wiley & Sons, Inc., New York.
- Tanabe, K. and M. Sagae (1992), “An exact Cholesky Decomposition and the generalized inverse of the variance-covariance matrix of the multinomial distribution, with applications”, *Journal of the Royal Statistical Society B*, **54**, 211–219.
- Trenkler, G. (2000), *On a Generalisation of the Covariance Matrix of the Multinomial Distribution*, Kluwer Academic Publishers, Dordrecht, chapter 4, 69–73.
- Vermeulen, L. A. (1967), “The solution of a certain polynomial equation”, *The Mathematical Gazette*, **51**, 308–309.
- Wansbeek, T. (1985), “Singuliere covariantiematrices en SUR-modellen: Enige opmerkingen over de specificatie van de Boer en Harkema”, *Kwantitatieve Methoden*, **18**, 99–102.

Watson, G. S. (1996), “Spectral decomposition of the covariance matrix of a multinomial”, *Journal of the Royal Statistical Society B*, **58**, 289–291.