Envelope Models and Methods

Dimension Reduction for Efficient Estimation in Multivariate Statistics

R. Dennis Cook
School of Statistics
University of Minnesota
Minneapolis, MN 55455, U.S.A.

October 23, 2013

Contents

Chapter 5

Envelope Algebra

We consider the linear algebra and statistical foundations of envelopes in this chapter. Section 5.1 contains a discussion of invariant and reducing subspaces, which are precursors to envelopes introduced in Section 5.2. Relationships between envelopes are discussed in Section 5.3. Most of the material in these sections comes from Cook *et al.* (2010).

5.1 Invariant and reducing subspaces

The following definition of invariant and reducing subspaces is a common construction in linear algebra (see, for example, Conway, 1990).

Definition 5.1 *A subspace $\mathcal{R}$ of $\mathbb{R}^r$ is an* invariant subspace *of $\mathbf{M} \in \mathbb{R}^{r\times r}$ if $\mathbf{M}\mathcal{R} \subseteq \mathcal{R}$; so $\mathbf{M}$ maps $\mathcal{R}$ to a subset of itself. $\mathcal{R}$ is a* reducing subspace *of $\mathbf{M}$ if $\mathcal{R}$ and $\mathcal{R}^\perp$ are both invariant subspaces of $\mathbf{M}$. If $\mathcal{R}$ is a reducing subspace of $\mathbf{M}$, we say that $\mathcal{R}$ reduces $\mathbf{M}$.*

The notion of reduction conveyed by this definition is incompatible with how reduction is usually understood in statistics. Nevertheless, it is a recognized term in linear algebra and functional analysis and it forms a convenient starting point for envelopes. It may not be straightforward to gain intuition about invariant and reducing subspaces from this definition alone. The following series of results provide characterizations that may facilitate intuition and will be helpful later.

The next lemma describes a matrix equation that characterizes invariant subspaces.

Lemma 5.1 *Let $\mathcal{R}$ be a u dimensional subspace of $\mathbb{R}^r$ and let $\mathbf{M} \in \mathbb{R}^{r\times r}$. Then $\mathcal{R}$ is an invariant subspace of $\mathbf{M}$ if and only if, for any $\mathbf{A} \in \mathbb{R}^{r\times s}$, $s \geq u$, such that $\mathrm{span}(\mathbf{A}) = \mathcal{R}$, there exists a $\mathbf{B} \in \mathbb{R}^{s\times s}$ such that $\mathbf{MA} = \mathbf{AB}$.*

PROOF. Suppose there is a $\mathbf{B}$ that satisfies $\mathbf{MA} = \mathbf{AB}$. For every $\mathbf{v} \in \mathcal{R}$ there is a $\mathbf{t} \in \mathbb{R}^s$ so that $\mathbf{v} = \mathbf{At}$. Consequently, $\mathbf{Mv} = \mathbf{MAt} = \mathbf{ABt} \in \mathcal{R}$, which implies that $\mathcal{R}$ is an invariant subspace of $\mathbf{M}$.

Suppose that $\mathcal{R}$ is an invariant subspace of $\mathbf{M}$, and let $\mathbf{a}_j$, $j = 1, \ldots, s$ denote the columns of $\mathbf{A}$. Then $\mathbf{Ma}_j \in \mathcal{R}$, $j = 1, \ldots, s$. Consequently, $\mathrm{span}(\mathbf{MA}) \subseteq \mathcal{R}$, which implies there is a $\mathbf{B} \in \mathbb{R}^{s\times s}$ such that $\mathbf{MA} = \mathbf{AB}$. □

The next result gives a connection between reducing subspaces and the eigenvectors of a symmetric $\mathbf{M}$. It does not require that the eigenvalues for $\mathbf{M}$ be unique.

Lemma 5.2 *Suppose that $\mathcal{R}$ reduces $\mathbf{M} \in \mathbb{S}^{r\times r}$. Then $\mathbf{M}$ has a spectral decomposition with eigenvectors in $\mathcal{R}$ or in $\mathcal{R}^\perp$.*

PROOF. Let $\mathbf{A}_0 \in \mathbb{R}^{r\times u}$ be a semi-orthogonal matrix whose columns span $\mathcal{R}$ and let $\mathbf{A}_1$ be its completion, such that $(\mathbf{A}_0, \mathbf{A}_1) \equiv \mathbf{A}$ is an orthogonal matrix. Because $\mathbf{M}\mathcal{R} \subseteq \mathcal{R}$ and $\mathbf{M}\mathcal{R}^\perp \subseteq \mathcal{R}^\perp$, it follows from Lemma 5.1 there exist matrices $\mathbf{B}_0 \in \mathbb{R}^{u\times u}$ and $\mathbf{B}_1 \in \mathbb{R}^{(r-u)\times(r-u)}$ such that $\mathbf{MA}_0 = \mathbf{A}_0\mathbf{B}_0$ and $\mathbf{MA}_1 = \mathbf{A}_1\mathbf{B}_1$. Hence

$$\mathbf{M}\begin{pmatrix}\mathbf{A}_0 & \mathbf{A}_1\end{pmatrix} = \begin{pmatrix}\mathbf{A}_0 & \mathbf{A}_1\end{pmatrix}\begin{pmatrix}\mathbf{B}_0 & 0\\ 0 & \mathbf{B}_1\end{pmatrix} \Leftrightarrow \mathbf{M} = \mathbf{A}\begin{pmatrix}\mathbf{B}_0 & 0\\ 0 & \mathbf{B}_1\end{pmatrix}\mathbf{A}^T.$$

Because $\mathbf{M}$ is symmetric, so must be $\mathbf{B}_0$ and $\mathbf{B}_1$. Hence $\mathbf{B}_0$ and $\mathbf{B}_1$ have spectral decompositions $\mathbf{C}_0\boldsymbol{\Lambda}_0\mathbf{C}_0^T$ and $\mathbf{C}_1\boldsymbol{\Lambda}_1\mathbf{C}_1^T$ for some diagonal matrices $\boldsymbol{\Lambda}_0$ and $\boldsymbol{\Lambda}_1$ and orthogonal matrices $\mathbf{C}_0$ and $\mathbf{C}_1$. Let $\mathbf{C} = \mathrm{diag}(\mathbf{C}_0, \mathbf{C}_1)$ and $\boldsymbol{\Lambda} = \mathrm{diag}(\boldsymbol{\Lambda}_0, \boldsymbol{\Lambda}_1)$. Then,

$$\mathbf{M} = \mathbf{AC}\boldsymbol{\Lambda}\mathbf{C}^T\mathbf{A}^T \equiv \mathbf{D}\boldsymbol{\Lambda}\mathbf{D}^T, \tag{5.1}$$

where $\mathbf{D} = \mathbf{AC}$. The first u columns of $\mathbf{D}$, which form the matrix $\mathbf{A}_0\mathbf{C}_0$, span $\mathcal{R}$. Moreover, $\mathbf{D}$ is an orthogonal matrix, and thus (5.1) is a spectral decomposition of $\mathbf{M}$ with eigenvectors in $\mathcal{R}$ or $\mathcal{R}^\perp$. □

The next proposition gives a characterization of $\mathbf{M}$ in terms of its projections onto its reducing subspaces. It shows in effect that $\mathbf{M}$ can be decomposed into a sum of orthogonal matrices, and it was used in forming the parameters in the envelope models of Sections 1.2 and 2.3 (cf. (2.10)). Representation (5.2) could be used as the definition of a reducing subspace in the context of the proposition.

Proposition 5.1 *$\mathcal{R}$ reduces $\mathbf{M} \in \mathbb{R}^{r\times r}$ if and only if $\mathbf{M}$ can be written in the form*

$$\mathbf{M} = \mathbf{P}_{\mathcal{R}}\mathbf{M}\mathbf{P}_{\mathcal{R}} + \mathbf{Q}_{\mathcal{R}}\mathbf{M}\mathbf{Q}_{\mathcal{R}}. \tag{5.2}$$

PROOF. Assume that $\mathbf{M}$ can be written as in (5.2). Then for any $\mathbf{v} \in \mathcal{R}$, $\mathbf{M}\mathbf{v} \in \mathcal{R}$, and for $\mathbf{v} \in \mathcal{R}^{\perp}$, $\mathbf{M}\mathbf{v} \in \mathcal{R}^{\perp}$. Consequently, $\mathcal{R}$ reduces $\mathbf{M}$.

Next, assume that $\mathcal{R}$ reduces $\mathbf{M}$. We must show that $\mathbf{M}$ satisfies (5.2). Let $u = \dim(\mathcal{R})$. It follows from Lemma 5.1 that there is a $\mathbf{B} \in \mathbb{R}^{u\times u}$ that satisfies $\mathbf{M}\mathbf{A} = \mathbf{A}\mathbf{B}$, where $\mathbf{A} \in \mathbb{R}^{r\times u}$ and $\mathrm{span}(\mathbf{A}) = \mathcal{R}$. This implies $\mathbf{Q}_{\mathcal{R}}\mathbf{M}\mathbf{A} = 0$ which is equivalent to $\mathbf{Q}_{\mathcal{R}}\mathbf{M}\mathbf{P}_{\mathcal{R}} = 0$. By the same logic applied to $\mathcal{R}^{\perp}$, $\mathbf{P}_{\mathcal{R}}\mathbf{M}\mathbf{Q}_{\mathcal{R}} = 0$. Consequently,

$$\mathbf{M} = (\mathbf{P}_{\mathcal{R}} + \mathbf{Q}_{\mathcal{R}})\mathbf{M}(\mathbf{P}_{\mathcal{R}} + \mathbf{Q}_{\mathcal{R}}) = \mathbf{P}_{\mathcal{R}}\mathbf{M}\mathbf{P}_{\mathcal{R}} + \mathbf{Q}_{\mathcal{R}}\mathbf{M}\mathbf{Q}_{\mathcal{R}}.$$

$\square$

An eigenspace of $\mathbf{M} \in \mathbb{S}^{r\times r}$ is a subspace of $\mathbb{R}^r$ spanned by the eigenvectors of $\mathbf{M}$ with the same eigenvalue. The dimension of an eigenspace is then equal to the multiplicity of the corresponding eigenvector. An eigenspace of $\mathbf{M}$ reduces $\mathbf{M}$. However, a reducing subspace of $\mathbf{M}$ need not be an eigenspace of $\mathbf{M}$. For example, if $\mathbf{M}$ has an eigenvalue λ_1 with multiplicity 1 and an eigenvalue λ_2 with multiplicity 2, then $\mathbf{M}$ is reduced by the span of the eigenvector associated λ_1 and any vector in the eigenspace of λ_2, but this reducing subspace is not an eigenspace. The span of any vector in the eigenspace of λ_2 also reduces $\mathbf{M}$ but is not an eigenspace of $\mathbf{M}$.

Corollary 5.1 describes consequences of Proposition 5.1, including a relationship between reducing subspaces of $\mathbf{M}$ and $\mathbf{M}^{-1}$, when $\mathbf{M}$ is non-singular. Results 3 and 4 in this corollary were used during the analysis of the log likelihood L in Section 2.4. Connecting the notation of that section with the corollary, $\mathbf{\Omega} = \mathbf{A}^T\mathbf{M}\mathbf{A}$ and $\mathbf{\Omega}_0 = \mathbf{A}_0\mathbf{M}\mathbf{A}_0$.

Corollary 5.1 *Let $\mathcal{R}$ reduce $\mathbf{M} \in \mathbb{R}^{r\times r}$, let $\mathbf{A} \in \mathbb{R}^{r\times u}$ be a semi-orthogonal basis matrix for $\mathcal{R}$, and let $\mathbf{A}_0$ be a semi-orthogonal basis matrix for $\mathcal{R}^{\perp}$. Then*

1. *$\mathbf{M}$ and $\mathbf{P}_{\mathcal{R}}$, and $\mathbf{M}$ and $\mathbf{Q}_{\mathcal{R}}$ commute.*
2. *$\mathcal{R} \subseteq \mathrm{span}(\mathbf{M})$ if and only if $\mathbf{A}^T\mathbf{M}\mathbf{A}$ is full rank.*
3. *$|\mathbf{M}| = |\mathbf{A}^T\mathbf{M}\mathbf{A}| \times |\mathbf{A}_0^T\mathbf{M}\mathbf{A}_0|$.*
4. *If $\mathbf{M}$ is full rank then*

$$\mathbf{M}^{-1} = \mathbf{A}(\mathbf{A}^T\mathbf{M}\mathbf{A})^{-1}\mathbf{A}^T + \mathbf{A}_0(\mathbf{A}_0^T\mathbf{M}\mathbf{A}_0)^{-1}\mathbf{A}_0^T. \tag{5.3}$$

5. *If* $\mathcal{R} \subseteq \mathrm{span}(\mathbf{M})$ *then*

$$\mathbf{M}^{\dagger} = \mathbf{A}(\mathbf{A}^T\mathbf{M}\mathbf{A})^{-1}\mathbf{A}^T + \mathbf{A}_0(\mathbf{A}_0^T\mathbf{M}\mathbf{A}_0)^{\dagger}\mathbf{A}_0^T.$$

PROOF. The first conclusion follows immediately from Proposition 5.1.

To show the second conclusion, first assume that $\mathbf{A}^T\mathbf{M}\mathbf{A}$ is full rank. Then, from Lemma 5.1, $\mathbf{B}$ must be full rank in the representation $\mathbf{M}\mathbf{A} = \mathbf{A}\mathbf{B}$. Consequently, any vector in $\mathcal{R}$ can be written as a linear combination of the columns of $\mathbf{M}$ and thus $\mathcal{R} \subseteq \mathrm{span}(\mathbf{M})$. Next, assume that $\mathcal{R} \subseteq \mathrm{span}(\mathbf{M})$. Then there is a full rank matrix $\mathbf{V} \in \mathbb{R}^{r \times u}$ such that $\mathbf{M}\mathbf{V} = \mathbf{A}$ and thus that $\mathbf{A}^T\mathbf{M}\mathbf{V} = \mathbf{I}_u$. Substituting $\mathbf{M}$ from Proposition 5.1, we have $(\mathbf{A}^T\mathbf{M}\mathbf{A})(\mathbf{A}^T\mathbf{V}) = \mathbf{I}_u$. It follows that $\mathbf{A}^T\mathbf{M}\mathbf{A}$ is of full rank.

For the fourth conclusion, since $\mathbf{M}$ is full rank $\mathcal{R} \subseteq \mathrm{span}(\mathbf{M})$ and $\mathcal{R}^{\perp} \subseteq \mathrm{span}(\mathbf{M})$. Consequently, both $\mathbf{A}^T\mathbf{M}\mathbf{A}$ and $\mathbf{A}_0^T\mathbf{M}\mathbf{A}_0$ are full rank. Thus the right hand side of (5.3) is defined. Meanwhile, note that $\mathbf{P}_{\mathcal{R}} = \mathbf{A}\mathbf{A}^T$ and $\mathbf{Q}_{\mathcal{R}} = \mathbf{A}_0\mathbf{A}_0^T$. Hence, by (5.2), $\mathbf{M} = \mathbf{A}\mathbf{A}^T\mathbf{M}\mathbf{A}\mathbf{A}^T + \mathbf{A}_0\mathbf{A}_0^T\mathbf{M}\mathbf{A}_0\mathbf{A}_0^T$. Multiply this and the right hand side of (5.3) to complete the proof. The final conclusion follows similarly: Since $\mathcal{R} \subseteq \mathrm{span}(\mathbf{M})$, $\mathbf{A}^T\mathbf{M}\mathbf{A}$ is full rank. The conclusion follows by straightforwardly checking the conditions for the Moore-Penrose inverse. □

Lemma 5.3 *Let* $\mathcal{R}$ *reduce* $\mathbf{M} \in \mathbb{R}^{r \times r}$. *Then* $\mathbf{M}\mathcal{R} = \mathcal{R}$ *if and only if* $\mathcal{R} \subseteq \mathrm{span}(\mathbf{M})$.

PROOF. Assume that $\mathbf{M}\mathcal{R} = \mathcal{R}$. Then, with $\mathbf{A}$ as defined in Lemma 5.1 and $s = u$, $\mathbf{M}\mathbf{A} = \mathbf{A}\mathbf{B}$ for some full rank matrix $\mathbf{B} \in \mathbb{R}^{u \times u}$. Consequently, $\mathbf{A}^T\mathbf{M}\mathbf{A}$ is full rank. It follows from Corollary 5.1 that $\mathcal{R} \subseteq \mathrm{span}(\mathbf{M})$.

Assume that $\mathcal{R} \subseteq \mathrm{span}(\mathbf{M})$. Then it follows from Corollary 5.1 that $\mathbf{A}^T\mathbf{M}\mathbf{A}$ is of full rank. Thus, $\mathbf{B}$ must have full rank in the representation $\mathbf{M}\mathbf{A} = \mathbf{A}\mathbf{B}$, which implies $\mathbf{M}\mathcal{R} = \mathcal{R}$. □

Lemma 5.3 gives necessary and sufficient conditions for $\mathbf{M}$ to map $\mathcal{R}$ to itself, rather than to a proper subset of $\mathcal{R}$. We will nearly always use reducing subspaces in contexts where $\mathbf{M}$ is a full rank covariance matrix. Lemma 5.3 holds trivially in such cases since then it is always true that $\mathcal{R} \subseteq \mathrm{span}(\mathbf{M}) = \mathbb{R}^r$.

As mentioned in connection with Lemma 5.2, there is a relationship between the eigenstructure of a symmetric matrix $\mathbf{M}$ and its reducing subspaces. By definition, any invariant subspace of $\mathbf{M} \in \mathbb{S}^{r\times r}$ is also a reducing subspace of $\mathbf{M}$. In particular, it follows from Proposition 5.1 that the subspace spanned by any set of eigenvectors of $\mathbf{M}$ is a reducing subspace of $\mathbf{M}$. This connection is formalized as follows.

Proposition 5.2 *Let $\mathcal{R}$ be a subspace of $\mathbb{R}^r$ and let $\mathbf{M} \in \mathbb{S}^{r\times r}$. Assume that $\mathbf{M}$ has $q \le r$ distinct eigenvalues, and let $\mathbf{P}_i$, $i = 1,\ldots,q$, indicate the projections on the corresponding eigenspaces. Then the following statements are equivalent:*

1. *$\mathcal{R}$ reduces $\mathbf{M}$,*
2. *$\mathcal{R} = \sum_{i=1}^q \mathbf{P}_i\mathcal{R}$,*
3. *$\mathbf{P}_\mathcal{R} = \sum_{i=1}^q \mathbf{P}_i\mathbf{P}_\mathcal{R}\mathbf{P}_i$,*
4. *$\mathbf{M}$ and $\mathbf{P}_\mathcal{R}$ commute.*

PROOF. **equivalence of 1 and 4:** If $\mathcal{R}$ reduces $\mathbf{M}$, then it follows immediately from Corollary 5.1 that $\mathbf{M}$ and $\mathbf{P}_\mathcal{R}$ commute. If $\mathbf{M}$ and $\mathbf{P}_\mathcal{R}$ commute, then $\mathbf{P}_\mathcal{R}\mathbf{M}\mathbf{Q}_\mathcal{R} = \mathbf{M}\mathbf{P}_\mathcal{R}\mathbf{Q}_\mathcal{R} = 0$, and thus it follows from Proposition 5.1 that $\mathcal{R}$ reduces $\mathbf{M}$.

1 implies 2: If $\mathbf{v} \in \mathcal{R}$, then

$$\mathbf{v} = \mathbf{I}_p\mathbf{v} = \left(\sum_{i=1}^q \mathbf{P}_i\right)\mathbf{v} = \sum_{i=1}^q \mathbf{P}_i\mathbf{v} \in \sum_{i=1}^q \mathbf{P}_i\mathcal{R}.$$

Hence $\mathcal{R} \subseteq \sum_{i=1}^q \mathbf{P}_i\mathcal{R}$. Conversely, if $\mathbf{v} \in \sum_{i=1}^q \mathbf{P}_i\mathcal{R}$, then $\mathbf{v}$ can be written as a linear combination of $\mathbf{P}_1\mathbf{v}_1,\ldots,\mathbf{P}_q\mathbf{v}_q$ where $\mathbf{v}_1,\ldots,\mathbf{v}_q$ belong to $\mathcal{R}$. By Lemma 5.2, $\mathbf{P}_i\mathbf{w} \in \mathcal{R}$ for any $\mathbf{w} \in \mathcal{R}$. Hence any linear combination of $\mathbf{P}_1\mathbf{v}_1,\ldots,\mathbf{P}_q\mathbf{v}_q$, with $\mathbf{v}_1,\ldots,\mathbf{v}_q$ belonging to $\mathcal{R}$, belongs to $\mathcal{R}$. That is, $\sum_{i=1}^q \mathbf{P}_i\mathcal{R} \subseteq \mathcal{R}$.

2 implies 3: If $\mathbf{v} \in \mathcal{R}$ then, from the previous step, $\mathbf{P}_i\mathbf{v} \in \mathcal{R}$, $i = 1,\ldots,q$. Hence

$$\left(\sum_{i=1}^q \mathbf{P}_i\mathbf{P}_\mathcal{R}\mathbf{P}_i\right)\mathbf{v} = \sum_{i=1}^q \mathbf{P}_i\mathbf{v} = \mathbf{v} = \mathbf{P}_\mathcal{R}\mathbf{v}.$$

Now let $\mathbf{v} \in \mathcal{R}^\perp$. Then, $\mathbf{v} \perp \mathbf{P}_i\mathcal{R}$ for each i. Because $\mathbf{P}_i$ is self-adjoint we see that $\mathbf{P}_i\mathbf{v} \perp \mathcal{R}$ for each i. Consequently

$$\left(\sum_{i=1}^q \mathbf{P}_i\mathbf{P}_\mathcal{R}\mathbf{P}_i\right)\mathbf{v} = 0 = \mathbf{P}_\mathcal{R}\mathbf{v}.$$

It follows that $(\sum \mathbf{P}_i\mathbf{P}_\mathcal{R}\mathbf{P}_i)\mathbf{v} = \mathbf{P}_\mathcal{R}\mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^r$. Hence the two matrices are the same.

3 implies 1: Again, if $\mathbf{v} \in \mathcal{R}$ then $\mathbf{P}_i\mathbf{v} \in \mathcal{R}$, $i = 1, \dots, q$. Hence, indicating with m_i, $i = 1, \dots, q$ the distinct eigenvalues of M we have

$$\mathbf{P}_{\mathcal{R}}\mathbf{M}\mathbf{v} = \sum_{i=1}^{q} m_i \mathbf{P}_i \mathbf{P}_{\mathcal{R}} \mathbf{P}_i \mathbf{P}_{\mathcal{R}} \mathbf{P}_i \mathbf{v} = \sum_{i=1}^{q} m_i \mathbf{P}_i \mathbf{P}_{\mathcal{R}} \mathbf{P}_i \mathbf{v} = \mathbf{M}\mathbf{v}.$$

It follows that $\mathbf{M}\mathcal{R} \subseteq \mathcal{R}$. □

Lemma 5.4 *The intersection of any two reducing subspaces of* $\mathbf{M} \in \mathbb{R}^{r \times r}$ *is also a reducing subspace of* $\mathbf{M}$.

PROOF. Let $\mathcal{R}_1$ and $\mathcal{R}_2$ be reducing subspaces of $\mathbf{M}$. Then by definition $\mathbf{M}\mathcal{R}_1 \subseteq \mathcal{R}_1$ and $\mathbf{M}\mathcal{R}_2 \subseteq \mathcal{R}_2$. Clearly, if $\mathbf{v} \in \mathcal{R}_1 \cap \mathcal{R}_2$ then $\mathbf{M}\mathbf{v} \in \mathcal{R}_1 \cap \mathcal{R}_2$, and it follows that the intersection is an invariant subspace of $\mathbf{M}$. The same argument shows that if $\mathbf{v} \in (\mathcal{R}_1 \cap \mathcal{R}_2)^{\perp} = \mathcal{R}_1^{\perp} + \mathcal{R}_2^{\perp}$ then $\mathbf{M}\mathbf{v} \in \mathcal{R}_1^{\perp} + \mathcal{R}_2^{\perp}$: If $\mathbf{v} \in \mathcal{R}_1^{\perp} + \mathcal{R}_2^{\perp}$ then it can be written as $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$, where $\mathbf{v} \in \mathcal{R}_1^{\perp}$ and $\mathbf{v} \in \mathcal{R}_2^{\perp}$. Then $\mathbf{M}\mathbf{v} = \mathbf{M}\mathbf{v}_1 + \mathbf{M}\mathbf{v}_2 \in \mathcal{R}_1^{\perp} + \mathcal{R}_2^{\perp}$.
□

Lemma 5.5 *Let* $\mathbf{M}_j \in \mathbb{R}^{r_j \times r_j}$*, and let* $\mathcal{R} \subseteq \mathbb{R}^{r_j}$*,* $j = 1, 2$*. Then* $\mathcal{R} = \mathcal{R}_1 \otimes \mathcal{R}_2 := \{\mathbf{v}_1 \otimes \mathbf{v}_2 | \mathbf{v}_1 \in \mathcal{R}_1, \mathbf{v}_2 \in \mathcal{R}_2\}$ *reduces* $\mathbf{M} = \mathbf{M}_1 \otimes \mathbf{M}_2$ *if and only if* $\mathcal{R}_j$ *reduce* $\mathbf{M}_j$*,* $j = 1, 2$.

PROOF.

If $\mathcal{R}_j$ reduce $\mathbf{M}_j$, $j = 1, 2$. Then

$$\begin{aligned}
\mathbf{M} &= (\mathbf{P}_{\mathcal{R}_1}\mathbf{M}_1\mathbf{P}_{\mathcal{R}_1} + \mathbf{Q}_{\mathcal{R}_1}\mathbf{M}_1\mathbf{Q}_{\mathcal{R}_1}) \otimes (\mathbf{P}_{\mathcal{R}_2}\mathbf{M}_2\mathbf{P}_{\mathcal{R}_2} + \mathbf{Q}_{\mathcal{R}_2}\mathbf{M}_1\mathbf{Q}_{\mathcal{R}_2}) \\
&= \mathbf{P}_{\mathcal{R}_1}\mathbf{M}_1\mathbf{P}_{\mathcal{R}_1} \otimes \mathbf{P}_{\mathcal{R}_2}\mathbf{M}_2\mathbf{P}_{\mathcal{R}_2} + \mathbf{P}_{\mathcal{R}_1}\mathbf{M}_1\mathbf{P}_{\mathcal{R}_1} \otimes \mathbf{Q}_{\mathcal{R}_2}\mathbf{M}_2\mathbf{Q}_{\mathcal{R}_2} \\
&\quad + \mathbf{Q}_{\mathcal{R}_1}\mathbf{M}_1\mathbf{Q}_{\mathcal{R}_1} \otimes \mathbf{P}_{\mathcal{R}_2}\mathbf{M}_2\mathbf{P}_{\mathcal{R}_2} + \mathbf{Q}_{\mathcal{R}_1}\mathbf{M}_1\mathbf{Q}_{\mathcal{R}_1} \otimes \mathbf{Q}_{\mathcal{R}_2}\mathbf{M}_2\mathbf{Q}_{\mathcal{R}_2} \\
&= (\mathbf{P}_{\mathcal{R}_1} \otimes \mathbf{P}_{\mathcal{R}_2})\mathbf{M}(\mathbf{P}_{\mathcal{R}_1} \otimes \mathbf{P}_{\mathcal{R}_2}) + (\mathbf{P}_{\mathcal{R}_1} \otimes \mathbf{Q}_{\mathcal{R}_2})\mathbf{M}(\mathbf{P}_{\mathcal{R}_1} \otimes \mathbf{Q}_{\mathcal{R}_2}) \\
&\quad + (\mathbf{Q}_{\mathcal{R}_1} \otimes \mathbf{P}_{\mathcal{R}_2})\mathbf{M}(\mathbf{Q}_{\mathcal{R}_1} \otimes \mathbf{P}_{\mathcal{R}_2}) + (\mathbf{Q}_{\mathcal{R}_1} \otimes \mathbf{Q}_{\mathcal{R}_2})\mathbf{M}(\mathbf{Q}_{\mathcal{R}_1} \otimes \mathbf{Q}_{\mathcal{R}_2} \\
&= (\mathbf{P}_{\mathcal{R}_1} \otimes \mathbf{P}_{\mathcal{R}_2})\mathbf{M}(\mathbf{P}_{\mathcal{R}_1} \otimes \mathbf{P}_{\mathcal{R}_2}) + (\mathbf{I}_{r_1 r_2} - \mathbf{P}_{\mathcal{R}_1} \otimes \mathbf{P}_{\mathcal{R}_2})\mathbf{M}(\mathbf{I}_{r_1 r_2} - \mathbf{P}_{\mathcal{R}_1} \otimes \mathbf{P}_{\mathcal{R}_2}),
\end{aligned}$$

where the last step follows because $\mathcal{R}_j$ reduces $\mathbf{M}_j$, $j = 1, 2$.

If $\mathcal{R}$ reduces $\mathbf{M}$ then by Proposition 5.1 we have $\mathbf{P}_{\mathcal{R}}\mathbf{M}\mathbf{Q}_{\mathcal{R}} = 0$. Since $\mathbf{P}_{\mathcal{R}} = \mathbf{P}_{\mathcal{R}_1} \otimes \mathbf{P}_{\mathcal{R}_2}$,

$$\begin{aligned}\mathbf{P}_{\mathcal{R}}\mathbf{M}\mathbf{Q}_{\mathcal{R}} &= (\mathbf{P}_{\mathcal{R}_1} \otimes \mathbf{P}_{\mathcal{R}_2})(\mathbf{M}_1 \otimes \mathbf{M}_2)(\mathbf{I}_{r_1 r_2} - \mathbf{P}_{\mathcal{R}_1} \otimes \mathbf{P}_{\mathcal{R}_2}) \\ &= \mathbf{P}_{\mathcal{R}_1}\mathbf{M}_1 \otimes \mathbf{P}_{\mathcal{R}_2}\mathbf{M}_2 - \mathbf{P}_{\mathcal{R}_1}\mathbf{M}_1\mathbf{P}_{\mathcal{R}_1} \otimes \mathbf{P}_{\mathcal{R}_2}\mathbf{M}_2\mathbf{P}_{\mathcal{R}_2}.\end{aligned}$$

Now, $\mathbf{P}_{\mathcal{R}}\mathbf{M}\mathbf{Q}_{\mathcal{R}} = 0$ if and only if $\mathbf{P}_{\mathcal{R}_j}\mathbf{M}_j = \mathbf{P}_{\mathcal{R}_j}\mathbf{M}_j\mathbf{P}_{\mathcal{R}_j}$, $j = 1, 2$. Equivalently, we must have $\mathbf{P}_{\mathcal{R}_j}\mathbf{M}_j\mathbf{Q}_{\mathcal{R}_j} = 0$, $j = 1, 2$. The conclusion follows from Proposition 5.1. □

5.2 $\mathbf{M}$-Envelopes

Since the intersection of two reducing subspaces of $\mathbf{M} \in \mathbb{S}^{r \times r}$ is itself a reducing subspace of $\mathbf{M}$, it makes sense to talk about the smallest reducing subspace of $\mathbf{M}$ that contains a specified subspace $\mathcal{S}$.

Definition 5.2 *Let $\mathbf{M} \in \mathbb{S}^{r \times r}$ and let $\mathcal{S} \subseteq \mathrm{span}(\mathbf{M})$. The $\mathbf{M}$-envelope of $\mathcal{S}$, to be written as $\mathcal{E}_{\mathbf{M}}(\mathcal{S})$, is the intersection of all reducing subspaces of $\mathbf{M}$ that contain $\mathcal{S}$. The following convention will help avoid proliferation of notation: If $\mathbf{B} \in \mathbb{R}^{r \times d}$ and $\mathrm{span}(\mathbf{B}) = \mathcal{S}$ then $\mathcal{E}_{\mathbf{M}}(\mathbf{B}) := \mathcal{E}_{\mathbf{M}}(\mathrm{span}(\mathbf{B})) = \mathcal{E}_{\mathbf{M}}(\mathcal{S})$.*

This definition requires that $\mathcal{S} \subseteq \mathrm{span}(\mathbf{M})$. Since the column space of $\mathbf{M}$ is itself a reducing subspace of $\mathbf{M}$, this containment guarantees existence of the $\mathbf{M}$-envelope and it holds trivially if $\mathbf{M}$ is full rank, i.e, if $\mathrm{span}(\mathbf{M}) = \mathbb{R}^r$. Moreover, closure under intersection guarantees that the $\mathbf{M}$-envelope is in fact a reducing subspace of $\mathbf{M}$. Thus the $\mathbf{M}$-envelope of $\mathcal{S}$ can be interpreted as the unique smallest reducing subspace of $\mathbf{M}$ that contains $\mathcal{S}$, and represents a well-defined parameter in some statistical problems.

To aid intuition, consider the case where r eigenvalues of $\mathbf{M}$ are distinct. There are 2^r ways of dividing the eigenvectors of $\mathbf{M}$ into two sets and among them there is a unique way in which one of the two groups spans a subspace of minimal dimension that contains $\mathcal{S}$. This minimal subspace is $\mathcal{E}_{\mathbf{M}}(\mathcal{S})$. Thus, in this case, $\mathcal{E}_{\mathbf{M}}(\mathcal{S})$ is the smallest subspace that contains $\mathcal{S}$ and that is spanned by a subset of the eigenvectors of $\mathbf{M}$. The situation becomes more complicated if $\mathbf{M}$ has less than r distinct eigenvalues, which is why reducing subspaces were used in Definition 5.2.

Consider next the application of this definition in terms of Figure 2.1. We set $\mathbf{M} = \mathbf{\Sigma}$ and $\mathcal{S} = \mathcal{B} = \mathrm{span}(\boldsymbol{\beta})$. Then the envelope $\mathcal{E}_{\mathbf{\Sigma}}(\mathcal{B})$ shown in the figure is the smallest reducing subspace of $\mathbf{\Sigma}$ that contains $\mathcal{B}$. We know from Proposition 5.2 that the reducing subspaces of $\mathbf{\Sigma}$ correspond to eigenspaces of $\mathbf{\Sigma}$. Thus in the example we find the smallest eigenspace of $\mathbf{\Sigma}$ that contains $\mathcal{B}$. In the figure, $\mathcal{B}$ is equal to the eigenspace corresponding to the smaller eigenvalue of $\mathbf{\Sigma}$.

The following proposition, derived from Proposition 5.2 and Definition 5.2, gives a constructive characterization of $\mathbf{M}$-envelopes and provided a connection between $\mathcal{E}_{\mathbf{M}}(\mathcal{S})$ and the eigenspaces of $\mathbf{M}$.

Proposition 5.3 *Let $\mathbf{M} \in \mathbb{S}^{r\times r}$, let $\mathbf{P}_i$, $i = 1, \dots, q \le r$, be the projections onto the eigenspaces of $\mathbf{M}$, and let $\mathcal{S}$ be a subspace of $\operatorname{span}(\mathbf{M})$. Then $\mathcal{E}_{\mathbf{M}}(\mathcal{S}) = \sum_{i=1}^{q} \mathbf{P}_i\mathcal{S}$ and, as a special case, $\mathcal{E}_{\mathbf{I}_p}(\mathcal{S}) = \mathcal{S}$.*

PROOF. To prove that $\sum_{i=1}^{q} \mathbf{P}_i\mathcal{S}$ is the smallest reducing subspace of $\mathbf{M}$ that contains $\mathcal{S}$, it suffices to prove the following statements:

1. $\sum_{i=1}^{q} \mathbf{P}_i\mathcal{S}$ reduces $\mathbf{M}$.
2. $\mathcal{S} \subseteq \sum_{i=1}^{q} \mathbf{P}_i\mathcal{S}$.
3. If $\mathcal{T}$ reduces $\mathbf{M}$ and $\mathcal{S} \subseteq \mathcal{T}$, then $\sum_{i=1}^{q} \mathbf{P}_i\mathcal{S} \subseteq \mathcal{T}$.

Statement 1 follows from Proposition 5.2, as applied to $\mathcal{R} \equiv \sum_{i=1}^{q} \mathbf{P}_i\mathcal{S}$. Statement 2 holds because $\mathcal{S} = \{\mathbf{P}_1\mathbf{v} + \cdots + \mathbf{P}_q\mathbf{v} : \mathbf{v} \in \mathcal{S}\} \subseteq \sum_{i=1}^{q} \mathbf{P}_i\mathcal{S}$. Turning to statement 3, if $\mathcal{T}$ reduces $\mathbf{M}$, it can be written as $\mathcal{T} = \sum_{i=1}^{q} \mathbf{P}_i\mathcal{T}$ by Proposition 5.2. If, in addition, $\mathcal{S} \subseteq \mathcal{T}$ then we have $\mathbf{P}_i\mathcal{S} \subseteq \mathbf{P}_i\mathcal{T}$ for $i = 1, \dots, q$. Statement 3 follows since $\sum_{i=1}^{q} \mathbf{P}_i\mathcal{S} \subseteq \sum_{i=1}^{q} \mathbf{P}_i\mathcal{T} = \mathcal{T}$. □

5.3 Relationships between envelopes

In this section we consider relationships between various types of envelopes. We first investigate how the $\mathbf{M}$-envelope is modified by linear transformations of $\mathcal{S}$, keeping $\mathbf{M}$ fixed.

5.3.1 Invariance and equivariance

An envelope does not transform equivariantly for all linear transformations, but it does so for symmetric linear transformations that commute with $\mathbf{M}$:

Proposition 5.4 *Let* $\mathbf{K} \in \mathbb{S}^{r\times r}$ *commute with* $\mathbf{M} \in \mathbb{S}^{r\times r}$*, and let* $\mathcal{S}$ *be a subspace of* $\mathrm{span}(\mathbf{M})$*. Then* $\mathbf{K}\mathcal{S} \subseteq \mathrm{span}(\mathbf{M})$ *and*

$$\mathcal{E}_{\mathbf{M}}(\mathbf{K}\mathcal{S}) = \mathbf{K}\mathcal{E}_{\mathbf{M}}(\mathcal{S}). \tag{5.4}$$

If, in addition, $\mathcal{S} \subseteq \mathrm{span}(\mathbf{K})$ *and* $\mathcal{E}_{\mathbf{M}}(\mathcal{S})$ *reduces* $\mathbf{K}$*, then*

$$\mathcal{E}_{\mathbf{M}}(\mathbf{K}\mathcal{S}) = \mathcal{E}_{\mathbf{M}}(\mathcal{S}). \tag{5.5}$$

PROOF. Since $\mathcal{S} \subseteq \mathrm{span}(\mathbf{M})$ and $\mathbf{M}$ and $\mathbf{K}$ commute, we have

$$\mathbf{K}\mathcal{S} \subseteq \mathbf{K}\mathrm{span}(\mathbf{M}) = \mathrm{span}(\mathbf{K}\mathbf{M}) = \mathrm{span}(\mathbf{M}\mathbf{K}) \subseteq \mathrm{span}(\mathbf{M}).$$

Turning to (5.4), because $\mathbf{K}$ and $\mathbf{M}$ commute, they can be diagonalized simultaneously by an orthogonal matrix, say $\mathbf{U}$. Recall that $\mathbf{P}_i$ is the projection on the i-th eigenspace of $\mathbf{M}$, and let $d_i = \mathrm{rank}(\mathbf{P}_i)$. Partition $\mathbf{U} = (\mathbf{U}_1, \ldots, \mathbf{U}_q)$, where $\mathbf{U}_i$ contains d_i columns, $i = 1, \ldots, q$. Without loss of generality, we can assume that $\mathbf{U}_i\mathbf{U}_i^T = \mathbf{P}_i$ for $i = 1, \ldots, q$. Then $\mathbf{K}$ can be written as $\mathbf{U}_1\mathbf{\Lambda}_1\mathbf{U}_1^T + \cdots \mathbf{U}_q\mathbf{\Lambda}_q\mathbf{U}_q^T$, where the $\mathbf{\Lambda}_i$'s are diagonal matrices of dimension $d_i \times d_i$. It follows that

$$\begin{aligned}\mathbf{K}\mathbf{P}_i =&(\mathbf{U}_1\mathbf{\Lambda}_1\mathbf{U}_1^T + \cdots + \mathbf{U}_q\mathbf{\Lambda}_q\mathbf{U}_q^T)\mathbf{U}_i\mathbf{U}_i^T\\ =&\mathbf{U}_i\mathbf{\Lambda}_i\mathbf{U}_i^T = \mathbf{U}_i\mathbf{U}_i^T(\mathbf{U}_1\mathbf{\Lambda}_1\mathbf{U}_1^T + \cdots \mathbf{U}_q\mathbf{\Lambda}_q\mathbf{U}_q^T) = \mathbf{P}_i\mathbf{K}.\end{aligned}$$

That is, $\mathbf{K}$ and $\mathbf{P}_i$ commute. Now, by Proposition 5.3, $\mathcal{E}_{\mathbf{M}}(\mathcal{S}) = \sum_i^q \mathbf{P}_i\mathcal{S}$. Hence

$$\begin{aligned}\mathbf{K}\mathcal{E}_{\mathbf{M}}(\mathcal{S}) =&\{\mathbf{K}\mathbf{P}_1\mathbf{h}_1 + \cdots + \mathbf{K}\mathbf{P}_q\mathbf{h}_q : \mathbf{h}_1, \ldots, \mathbf{h}_q \in \mathcal{S}\}\\ =&\{\mathbf{P}_1\mathbf{K}\mathbf{h}_1 + \cdots + \mathbf{P}_q\mathbf{K}\mathbf{h}_q : \mathbf{h}_1, \ldots, \mathbf{h}_q \in \mathcal{S}\} = \sum_{i=1}^q \mathbf{P}_i\mathbf{K}\mathcal{S}.\end{aligned}$$

By Proposition 5.3 again, the right hand side is $\mathcal{E}_{\mathbf{M}}(\mathbf{K}\mathcal{S})$.

Now suppose, in addition, that $\mathcal{S} \subseteq \mathrm{span}(\mathbf{K})$ and $\mathcal{E}_{\mathbf{M}}(\mathcal{S})$ reduces $\mathbf{K}$. Since $\mathbf{K}$ commutes with $\mathbf{M}$, $\mathrm{span}(\mathbf{K})$ is an invariant subspace of $\mathbf{M}$: for all $\mathbf{h} \in \mathbb{R}^r$, $\mathbf{M}\mathbf{K}\mathbf{h} = \mathbf{K}\mathbf{M}\mathbf{h} \subseteq \mathrm{span}(\mathbf{K})$. Since $\mathbf{M}$ is symmetric, we further have that $\mathrm{span}(\mathbf{K})$ reduces $\mathbf{M}$. In sum then, $\mathrm{span}(\mathbf{K})$ is a reducing subspace of $\mathbf{M}$ that contains $\mathcal{S}$ and thus it must contain the smallest reducing substance of $\mathbf{M}$ that contains $\mathcal{S}$; that is, $\mathcal{E}_{\mathbf{M}}(\mathcal{S}) \subseteq \mathrm{span}(\mathbf{K})$. By Lemma 5.3, then, $\mathbf{K}\mathcal{E}_{\mathbf{M}}(\mathcal{S}) = \mathcal{E}_{\mathbf{M}}(\mathcal{S})$, which, in conjunction with (5.4), implies (5.5). $\square$

To explore a useful consequence of Proposition 5.4, start with any function $f : \mathbb{R} \to \mathbb{R}$ with the properties $f(0) = 0$ and $f(x) \neq 0$ whenever $x \neq 0$, and then create $f^* : \mathbb{S}^{r\times r} \to \mathbb{S}^{r\times r}$ as follows. Let m_i and $\mathbf{P}_i$, $i = 1, \ldots, q$ indicate the distinct eigenvalues and the projections onto the corresponding eigenspaces for a matrix $\mathbf{M} \in \mathbb{S}^{r\times r}$, and let $f^*(\mathbf{M}) = \sum_{i=1}^{q} f(m_i)\mathbf{P}_i$. Then it is straightforward to verify that (i) $f^*(\mathbf{M})$ commutes with $\mathbf{M}$, (ii) any subspace $\mathcal{S} \subseteq \mathrm{span}(\mathbf{M})$ satisfies $\mathcal{S} \subseteq \mathrm{span}\{f^*(\mathbf{M})\}$, and (iii) $\mathcal{E}_{\mathbf{M}}(\mathcal{S})$ reduces $f^*(\mathbf{M})$. Hence, by Proposition 5.4 we have $\mathcal{E}_{\mathbf{M}}(f^*(\mathbf{M})\mathcal{S}) = \mathcal{E}_{\mathbf{M}}(\mathcal{S})$. Additionally, if $f(\cdot)$ is a strictly monotonic function then $\mathbf{M}$ and $f^*(\mathbf{M})$ have the same eigenspaces. It follows from Proposition 5.3 that $\mathcal{E}_{f^*(\mathbf{M})}(\mathcal{S}) = \mathcal{E}_{\mathbf{M}}(\mathcal{S})$. We summarize these results in the following corollary.

Corollary 5.2 *With f and f^* as previously defined, $\mathcal{E}_{\mathbf{M}}(f^*(\mathbf{M})\mathcal{S}) = \mathcal{E}_{\mathbf{M}}(\mathcal{S})$. If f is strictly monotonic then $\mathcal{E}_{f^*(\mathbf{M})}(\mathcal{S}) = \mathcal{E}_{\mathbf{M}}(\mathcal{S})$. In particular,*

$$\mathcal{E}_{\mathbf{M}}(\mathbf{M}^k\mathcal{S}) = \mathcal{E}_{\mathbf{M}}(\mathcal{S}) \textit{ for all } k \in \mathbb{R} \tag{5.6}$$

$$\mathcal{E}_{\mathbf{M}^k}(\mathcal{S}) = \mathcal{E}_{\mathbf{M}}(\mathcal{S}) \textit{ for all } k \in \mathbb{R} \textit{ with } k \neq 0. \tag{5.7}$$

We next consider settings in which envelopes are invariant under certain simultaneous changes in $\mathbf{M}$ and $\mathcal{S}$. The relationships established here will be useful later.

Let $\boldsymbol{\Delta} \in \mathbb{S}^{r\times r}$ be a positive definite matrix and let $\mathcal{S}$ be a u-dimensional subspace of $\mathbb{R}^r$. Let $\mathbf{G} \in \mathbb{R}^{r\times u}$ be a semi-orthogonal basis matrix for $\mathcal{S}$ and let $\mathbf{V} \in \mathbb{S}^{u\times u}$ be positive definite. Define $\boldsymbol{\Psi} = \boldsymbol{\Delta} + \mathbf{G}\mathbf{V}\mathbf{G}^T$. Then

Proposition 5.5 $\boldsymbol{\Delta}^{-1}\mathcal{S} = \boldsymbol{\Psi}^{-1}\mathcal{S}$ *and*

$$\mathcal{E}_{\boldsymbol{\Delta}}(\mathcal{S}) = \mathcal{E}_{\boldsymbol{\Psi}}(\mathcal{S}) = \mathcal{E}_{\boldsymbol{\Delta}}(\boldsymbol{\Delta}^{-1}\mathcal{S}) = \mathcal{E}_{\boldsymbol{\Psi}}(\boldsymbol{\Psi}^{-1}\mathcal{S}) = \mathcal{E}_{\boldsymbol{\Psi}}(\boldsymbol{\Delta}^{-1}\mathcal{S}) = \mathcal{E}_{\boldsymbol{\Delta}}(\boldsymbol{\Psi}^{-1}\mathcal{S}).$$

PROOF. Using a variant of the Woodbury identity for matrix inverses we have

$$\begin{aligned}\boldsymbol{\Psi}^{-1} &= \boldsymbol{\Delta}^{-1} - \boldsymbol{\Delta}^{-1}\mathbf{G}(\mathbf{V}^{-1} + \mathbf{G}^T\boldsymbol{\Delta}^{-1}\mathbf{G})^{-1}\mathbf{G}^T\boldsymbol{\Delta}^{-1},\\ \boldsymbol{\Delta}^{-1} &= \boldsymbol{\Psi}^{-1} - \boldsymbol{\Psi}^{-1}\mathbf{G}(-\mathbf{V}^{-1} + \mathbf{G}^T\boldsymbol{\Psi}^{-1}\mathbf{G})^{-1}\mathbf{G}^T\boldsymbol{\Psi}^{-1}.\end{aligned}$$

Multiplying both equations on the right by $\mathbf{G}$, The first implies $\mathrm{span}(\boldsymbol{\Psi}^{-1}\mathbf{G}) \subseteq \mathrm{span}(\boldsymbol{\Delta}^{-1}\mathbf{G})$; the second implies $\mathrm{span}(\boldsymbol{\Delta}^{-1}\mathbf{G}) \subseteq \mathrm{span}(\boldsymbol{\Psi}^{-1}\mathbf{G})$. Hence $\boldsymbol{\Psi}^{-1}\mathcal{S} = \boldsymbol{\Delta}^{-1}\mathcal{S}$. From this we have also that $\mathcal{E}_{\boldsymbol{\Psi}}(\boldsymbol{\Psi}^{-1}\mathcal{S}) = \mathcal{E}_{\boldsymbol{\Psi}}(\boldsymbol{\Delta}^{-1}\mathcal{S})$ and $\mathcal{E}_{\boldsymbol{\Delta}}(\boldsymbol{\Psi}^{-1}\mathcal{S}) = \mathcal{E}_{\boldsymbol{\Delta}}(\boldsymbol{\Delta}^{-1}\mathcal{S})$.

We next show that $\mathcal{E}_{\boldsymbol{\Delta}}(\mathcal{S}) = \mathcal{E}_{\boldsymbol{\Psi}}(\mathcal{S})$ by demonstrating that $\mathcal{R} \subseteq \mathbb{R}^r$ is a reducing subspace of $\boldsymbol{\Delta}$ that contains $\mathcal{S}$ if and only if it is a reducing subspace of $\boldsymbol{\Psi}$ that contains $\mathcal{S}$. Suppose $\mathcal{R}$ is a reducing subspace of $\boldsymbol{\Delta}$ that contains $\mathcal{S}$. Let $\boldsymbol{\alpha} \in \mathcal{R}$. Then

$\mathbf{\Psi}\boldsymbol{\alpha} = \mathbf{\Delta}\boldsymbol{\alpha} + \mathbf{G}\mathbf{V}\mathbf{G}^T\boldsymbol{\alpha}$. $\mathbf{\Delta}\boldsymbol{\alpha} \in \mathcal{R}$ because $\mathcal{R}$ reduces $\mathbf{\Delta}$; the second term on the right is a vector in $\mathcal{R}$ because $\mathcal{S} \subseteq \mathcal{R}$. Thus, $\mathcal{R}$ is a reducing subspace of $\mathbf{\Psi}$ and by construction it contains $\mathcal{S}$. Next, suppose $\mathcal{R}$ is a reducing subspace of $\mathbf{\Psi}$ that contains $\mathcal{S}$. The reverse implication follows similarly by reasoning in terms of $\mathbf{\Delta}\boldsymbol{\alpha} = \mathbf{\Psi}\boldsymbol{\alpha} - \mathbf{G}\mathbf{V}\mathbf{G}^T\boldsymbol{\alpha}$. We have $\mathbf{\Psi}\boldsymbol{\alpha} \in \mathcal{R}$ because $\mathcal{R}$ reduces $\mathbf{\Psi}$; the second term on the right is a vector in $\mathcal{R}$ because $\mathcal{S} \subseteq \mathcal{R}$. The remaining equalities follow immediately from (5.6). □

For a first application of Proposition 5.5, consider the multivariate linear regression model of Chapter 2 but now assume that the predictors $\mathbf{X}$ are random with covariance matrix $\mathbf{\Sigma}_{\mathbf{X}}$, so that $\mathbf{Y}$ and $\mathbf{X}$ have a joint distribution. The envelope that we wish to estimate is still $\mathcal{E}_{\mathbf{\Sigma}}(\mathcal{B})$ since the analysis is conditional on the observed predictors, but because the predictors are random we can express the marginal variance of $\mathbf{Y}$ as

$$\mathbf{\Sigma}_{\mathbf{Y}} = \mathbf{\Sigma} + \boldsymbol{\beta}\mathbf{\Sigma}_{\mathbf{X}}\boldsymbol{\beta}^T = \mathbf{\Sigma} + \mathbf{G}\mathbf{V}\mathbf{G}^T,$$

where $\mathbf{G}$ is a semi-orthogonal basis matrix for $\mathcal{B} = \mathrm{span}(\boldsymbol{\beta})$, $\boldsymbol{\beta} = \mathbf{G}\mathbf{A}$ and $\mathbf{V} = \mathbf{A}\mathbf{\Sigma}\mathbf{A}^T$ is positive definite since $\mathbf{A}$ must have full row rank. The second form for $\mathbf{\Sigma}$ matches the decomposition required for Proposition 5.5 with $\mathbf{\Psi} = \mathbf{\Sigma}_{\mathbf{Y}}$ and $\mathbf{\Delta} = \mathbf{\Sigma}$. Consequently, we have

$$\mathcal{E}_{\mathbf{\Sigma}}(\mathcal{B}) = \mathcal{E}_{\mathbf{\Sigma}_{\mathbf{Y}}}(\mathcal{B}) = \mathcal{E}_{\mathbf{\Sigma}}(\mathbf{\Sigma}^{-1}\mathcal{B}) = \mathcal{E}_{\mathbf{\Sigma}_{\mathbf{Y}}}(\mathbf{\Sigma}_{\mathbf{Y}}^{-1}\mathcal{B}) = \mathcal{E}_{\mathbf{\Sigma}_{\mathbf{Y}}}(\mathbf{\Sigma}^{-1}\mathcal{B}) = \mathcal{E}_{\mathbf{\Sigma}}(\mathbf{\Sigma}_{\mathbf{Y}}^{-1}\mathcal{B}).$$

5.3.2 Coordinate reduction

The next proposition shows a way to express an envelope in terms of reduced coordinates.

Proposition 5.6 *Let* $\mathbf{B}$ *be a semi-orthogonal basis matrix for a reducing subspace of* $\mathbf{M} \in \mathbb{S}^{r\times r}$ *that contains* $\mathcal{S} \subseteq \mathbb{R}^r$. *Then* $\mathcal{E}_{\mathbf{M}}(\mathcal{S}) = \mathbf{B}\mathcal{E}_{\mathbf{B}^T\mathbf{M}\mathbf{B}}(\mathbf{B}^T\mathcal{S})$.

PROOF. For the conclusion to be reasonable we need to have $\mathbf{B}^T\mathcal{S} \subseteq \mathrm{span}(\mathbf{B}^T\mathbf{M}\mathbf{B})$ (cf. Definition 5.2). To see that this requirement holds, let $\mathbf{A}$ be a basis matrix for $\mathcal{S}$. Then since $\mathcal{S} \subseteq \mathrm{span}(\mathbf{M})$ there is a matrix $\mathbf{C}$ so that

$$\mathbf{A} = \mathbf{M}\mathbf{C} = \mathbf{P}_{\mathbf{B}}\mathbf{M}\mathbf{P}_{\mathbf{B}}\mathbf{C} + \mathbf{Q}_{\mathbf{B}}\mathbf{M}\mathbf{Q}_{\mathbf{B}}\mathbf{C},$$

where the second equality follows from Proposition 5.1 since $\mathrm{span}(\mathbf{B})$ reduces $\mathbf{M}$. This implies that $\mathbf{B}^T\mathbf{A} = \mathbf{B}^T\mathbf{M}\mathbf{B}(\mathbf{B}^T\mathbf{C})$ from which the requirement follows.

Let $\mathcal{D} = \mathcal{E}_{\mathbf{B}^T\mathbf{MB}}(\mathbf{B}^T\mathcal{S})$ for notational convenience. The conclusion can be deduced from the following quantities:

$$\begin{aligned}
\mathbf{M} &= \mathbf{P_B M P_B} + \mathbf{Q_B M Q_B} \\
&= \mathbf{B}(\mathbf{B}^T\mathbf{MB})\mathbf{B}^T + \mathbf{Q_B M Q_B} \\
&= \mathbf{B}\left\{\mathbf{P}_{\mathcal{D}}(\mathbf{B}^T\mathbf{MB})\mathbf{P}_{\mathcal{D}} + \mathbf{Q}_{\mathcal{D}}(\mathbf{B}^T\mathbf{MB})\mathbf{Q}_{\mathcal{D}}\right\}\mathbf{B}^T + \mathbf{Q_B M Q_B} \\
&= (\mathbf{B}\mathbf{P}_{\mathcal{D}}\mathbf{B}^T)\mathbf{M}(\mathbf{B}\mathbf{P}_{\mathcal{D}}\mathbf{B}^T) + (\mathbf{B}\mathbf{Q}_{\mathcal{D}}\mathbf{B}^T + \mathbf{Q_B})\mathbf{M}(\mathbf{B}\mathbf{Q}_{\mathcal{D}}\mathbf{B}^T + \mathbf{Q_B}),
\end{aligned}$$

where the third equality follows from Proposition 5.1 because $\mathcal{D}$ reduces $\mathbf{B}^T\mathbf{MB}$. The final equation holds because $\mathbf{B}\mathbf{Q}_{\mathcal{D}}\mathbf{B}^T\mathbf{M}\mathbf{Q_B} = 0$ since $\mathrm{span}(\mathbf{B})$ reduces $\mathbf{M}$ and therefore $\mathbf{Q_B}$ and $\mathbf{M}$ commute. Since $\mathbf{B}\mathbf{P}_{\mathcal{D}}\mathbf{B}^T$ and $\mathbf{B}\mathbf{Q}_{\mathcal{D}}\mathbf{B}^T + \mathbf{Q_B}$ are orthogonal projections that sum to $\mathbf{I}_r$, it follows from Proposition 5.1 that $\mathbf{B}\mathbf{P}_{\mathcal{D}}\mathbf{B}^T$ reduces $\mathbf{M}$. Additionally, since $\mathbf{B}^T\mathcal{S} \subseteq \mathcal{D}$, we have $\mathcal{S} = \mathbf{BB}^T\mathcal{S} \subseteq \mathbf{B}\mathcal{D}$. Consequently,

$$\mathrm{span}(\mathbf{B}\mathbf{P}_{\mathcal{D}}\mathbf{B}^T) = \mathrm{span}(\mathbf{B}\mathbf{P}_{\mathcal{D}}) = \mathbf{B}\mathcal{D}$$

is a reducing subspace of $\mathbf{M}$ that contains $\mathcal{S}$. The equality

$$\mathcal{E}_{\mathbf{M}}(\mathcal{S}) = \mathbf{B}\mathcal{D} = \mathbf{B}\mathcal{E}_{\mathbf{B}^T\mathbf{MB}}(\mathbf{B}^T\mathcal{S})$$

follows from the minimality of $\mathcal{D}$. □

Proposition 5.6 says that we never need to deal with a singular $\mathbf{M} \in \mathbb{S}^{r\times r}$ because we can always choose $\mathbf{B}$ to be a semi-orthogonal basis matrix for $\mathrm{span}(\mathbf{M})$, construct the envelope $\mathcal{E}_{\mathbf{B}^T\mathbf{MB}}(\mathbf{B}^T\mathcal{S})$ in reduced coordinates and then transform back to the full coordinates. The non-singularity of $\mathbf{B}^T\mathbf{MB}$ follows from conclusion 2 of Corollary 5.1 because $\mathcal{S} \subseteq \mathrm{span}(\mathbf{M})$ by construction.

Proposition 5.6 posits information about a superset of $\mathcal{E}_{\mathbf{M}}(\mathcal{S})$. In contrast, the next proposition starts with a subset of $\mathcal{E}_{\mathbf{M}}(\mathcal{S})$. It will be used later as a basis of sequential algorithms for determining $\mathcal{E}_{\mathbf{M}}(\mathcal{S})$.

Proposition 5.7 *Let $\mathbf{B}$ be a semi-orthogonal basis matrix for a subspace $\mathcal{B}$ of $\mathcal{E}_{\mathbf{M}}(\mathcal{S})$, where $\mathbf{M} \in \mathbb{S}^{r\times r}$ and $\mathcal{S} \subseteq \mathrm{span}(\mathbf{M})$. Let $(\mathbf{B}, \mathbf{B}_0)$ be an orthogonal matrix. Then*

$$\mathbf{B}_0\mathcal{E}_{\mathbf{B}_0^T\mathbf{M}\mathbf{B}_0}(\mathbf{B}_0^T\mathcal{S}) \subseteq \mathcal{E}_{\mathbf{M}}(\mathcal{S}).$$

PROOF. Since $\mathcal{S} \subseteq \mathrm{span}(\mathbf{M})$ we have $\mathbf{B}_0^T\mathcal{S} \subseteq \mathrm{span}(\mathbf{B}_0^T\mathbf{M}\mathbf{B}_0)$, so the envelope $\mathcal{E}_{\mathbf{B}_0^T\mathbf{M}\mathbf{B}_0}(\mathbf{B}_0^T\mathcal{S})$ is well defined. The next step is to use Proposition 5.1 to show that $\mathbf{B}_0^T\mathcal{E}_{\mathbf{M}}(\mathcal{S})$ is a reducing subspace of $\mathbf{B}_0^T\mathbf{M}\mathbf{B}_0$ that contains $\mathbf{B}_0^T\mathcal{S}$.

Since $\mathcal{E}_{\mathbf{M}}(\mathcal{S})$ reduces $\mathbf{M}$, $\mathbf{M} = \mathbf{P}_{\mathcal{E}}\mathbf{M}\mathbf{P}_{\mathcal{E}} + \mathbf{Q}_{\mathcal{E}}\mathbf{M}\mathbf{Q}_{\mathcal{E}}$, and thus

$$\mathbf{B}_0^T\mathbf{M}\mathbf{B}_0 = \mathbf{B}_0^T\mathbf{P}_{\mathcal{E}}\mathbf{M}\mathbf{P}_{\mathcal{E}}\mathbf{B}_0 + \mathbf{B}_0^T\mathbf{Q}_{\mathcal{E}}\mathbf{M}\mathbf{Q}_{\mathcal{E}}\mathbf{B}_0.$$

Since $\mathcal{B} \subseteq \mathcal{E}_{\mathbf{M}}(\mathcal{S})$ we have the decomposition $\mathbf{P}_{\mathcal{E}} = \mathbf{P}_{\mathcal{B}} + \mathbf{P}_{\mathbf{Q}_{\mathcal{B}}\mathcal{E}}$ and thus

$$\mathbf{B}_0^T\mathbf{P}_{\mathcal{E}} = \mathbf{B}_0^T\mathbf{P}_{\mathbf{Q}_{\mathcal{B}}\mathcal{E}} = \mathbf{B}_0^T\mathbf{P}_{\mathbf{B}_0\mathbf{B}_0^T\mathcal{E}} = \mathbf{B}_0^T\mathbf{B}_0\mathbf{P}_{\mathbf{B}_0^T\mathcal{E}}\mathbf{B}_0^T = \mathbf{P}_{\mathbf{B}_0^T\mathcal{E}}\mathbf{B}_0^T.$$

Also, since $\mathcal{E}_{\mathbf{M}}^{\perp}(\mathcal{S}) \subseteq \mathcal{B}^{\perp}$, $\mathbf{Q}_{\mathcal{E}}\mathbf{Q}_{\mathcal{B}} = \mathbf{Q}_{\mathcal{E}}$. Substituting these relationships and $\mathbf{Q}_{\mathcal{B}} = \mathbf{B}_0\mathbf{B}_0^T$, we have

$$\mathbf{B}_0^T\mathbf{M}\mathbf{B}_0 = \mathbf{P}_{\mathbf{B}_0^T\mathcal{E}}(\mathbf{B}_0^T\mathbf{M}\mathbf{B}_0)\mathbf{P}_{\mathbf{B}_0^T\mathcal{E}} + \mathbf{B}_0^T\mathbf{Q}_{\mathcal{E}}\mathbf{B}_0(\mathbf{B}_0^T\mathbf{M}\mathbf{B}_0)\mathbf{B}_0^T\mathbf{Q}_{\mathcal{E}}\mathbf{B}_0,$$

where $\mathbf{P}_{\mathbf{B}_0^T\mathcal{E}}$ and $\mathbf{B}_0^T\mathbf{Q}_{\mathcal{E}}\mathbf{B}_0$ are orthogonal projections, as can be seen by direct calculation:

$$\mathbf{B}_0^T\mathbf{Q}_{\mathcal{E}}\mathbf{B}_0\mathbf{B}_0^T\mathbf{Q}_{\mathcal{E}}\mathbf{B}_0 = \mathbf{B}_0^T\mathbf{Q}_{\mathcal{E}}\mathbf{B}_0 \text{ and } \mathbf{B}_0^T\mathbf{Q}_{\mathcal{E}}\mathbf{B}_0\mathbf{B}_0^T\mathcal{E}_{\mathbf{M}}(\mathcal{S}) = \mathbf{B}_0^T\mathbf{Q}_{\mathcal{E}}\mathcal{E}_{\mathbf{M}}(\mathcal{S}) = 0.$$

Since $\mathrm{span}(\mathbf{B}_0^T\mathbf{P}_{\mathcal{E}})$ and $\mathrm{span}(\mathbf{B}_0^T\mathbf{Q}_{\mathcal{E}})$ are orthogonal subspaces, we can find their joint span as $\mathrm{span}(\mathbf{B}_0^T\mathbf{P}_{\mathcal{E}}) + \mathrm{span}(\mathbf{B}_0^T\mathbf{Q}_{\mathcal{E}}) = \mathrm{span}(\mathbf{B}_0^T)$. It follows that $\mathrm{span}(\mathbf{B}_0^T\mathbf{P}_{\mathcal{E}}) = \mathbf{B}_0^T\mathcal{E}_{\mathbf{M}}(\mathcal{S})$ is a reducing subspace of $\mathbf{B}_0^T\mathbf{M}\mathbf{B}_0$ that contains $\mathbf{B}_0^T\mathcal{S}$, since $\dim(\mathrm{span}(\mathbf{B}_0^T))$ is the same as the dimension of the matrix $\mathbf{B}_0^T\mathbf{M}\mathbf{B}_0$.

Since $\mathcal{E}_{\mathbf{B}_0^T\mathbf{M}\mathbf{B}_0}(\mathbf{B}_0^T\mathcal{S})$ is the smallest reducing subspace of $\mathbf{B}_0^T\mathbf{M}\mathbf{B}_0$ that contains $\mathbf{B}_0^T\mathcal{S}$, we have

$$\mathcal{E}_{\mathbf{B}_0^T\mathbf{M}\mathbf{B}_0}(\mathbf{B}_0^T\mathcal{S}) \subseteq \mathbf{B}_0^T\mathcal{E}_{\mathbf{M}}(\mathcal{S}). \tag{5.8}$$

Consequently, if $\mathbf{v} \in \mathcal{E}_{\mathbf{B}_0^T\mathbf{M}\mathbf{B}_0}(\mathbf{B}_0^T\mathcal{S})$ then $\mathbf{B}_0\mathbf{v} \in \mathbf{Q}_{\mathcal{B}}\mathcal{E}_{\mathbf{M}}(\mathcal{S}) \subseteq \mathcal{E}_{\mathbf{M}}(\mathcal{S})$, since $\mathcal{B} \subseteq \mathcal{E}_{\mathbf{M}}(\mathcal{S})$: letting $\mathbf{\Gamma}$ be a semi-orthogonal basis matrix for $\mathcal{E}_{\mathbf{M}}(\mathcal{S})$ and writing $\mathbf{B} = \mathbf{\Gamma}\mathbf{A}$, where $\mathbf{A}$ is a semi-orthogonal matrix, we have

$$\mathbf{Q}_{\mathcal{B}}\mathbf{\Gamma} = \mathbf{\Gamma} - \mathbf{P}_{\mathcal{B}}\mathbf{\Gamma} = \mathbf{\Gamma} - \mathbf{\Gamma}\mathbf{P}_{\mathbf{A}} = \mathbf{\Gamma}\mathbf{Q}_{\mathbf{A}}.$$

□