CHAPTER 6: SINGULAR-VALUE DECOMPOSITION (SVD)

6.1 Introduction

Having finished with the eigenvalue problem for $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is square, we now turn our attention to the general $N \times M$ case $\mathbf{Gm} = \mathbf{d}$. First, the eigenvalue problem, per se, does not exist for $\mathbf{Gm} = \mathbf{d}$ unless $N = M$. This is because $\mathbf{G}$ maps (transforms) a vector $\mathbf{m}$ from M-space into a vector $\mathbf{d}$ in N-space. The concept of "parallel" breaks down when the vectors lie in different dimensional spaces.

Since the eigenvalue problem is not defined for $\mathbf{G}$, we will try to construct a square matrix that includes $\mathbf{G}$ (and, as it will turn out, $\mathbf{G}^T$) for which the eigenvalue problem is defined. This eigenvalue problem will lead us to *singular-value decomposition* (*SVD*), a way to decompose $\mathbf{G}$ into the product of three matrices (two eigenvector matrices $\mathbf{V}$ and $\mathbf{U}$, associated with model and data spaces, respectively, and a singular-value matrix very similar to Λ from the eigenvalue problem for $\mathbf{A}$). Finally, it will lead us to the generalized inverse operator, defined in a way that is analogous to the inverse matrix to $\mathbf{A}$ found using eigenvalue/eigenvector analysis.

The end result of SVD is

$$\boxed{\begin{array}{cccc} \mathbf{G} = & \mathbf{U}_P & \mathbf{\Lambda}_P & \mathbf{V}_P^T \\ N\times M & N\times P & P\times P & P\times M \end{array}} \tag{6.1}$$

where $\mathbf{U}_P$ are the P N-dimensional eigenvectors of $\mathbf{GG}^T$, $\mathbf{V}_P$ are the P M-dimensional eigenvectors of $\mathbf{G}^T\mathbf{G}$, and $\mathbf{\Lambda}_P$ is the $P \times P$ diagonal matrix with P singular values (positive square roots of the nonzero eigenvalues shared by $\mathbf{GG}^T$ and $\mathbf{G}^T\mathbf{G}$) on the diagonal.

6.2 Formation of a New Matrix B

6.2.1 Formulating the Eigenvalue Problem With G

The way to construct an eigenvalue problem that includes $\mathbf{G}$ is to form a square $(N + M) \times (N + M)$ matrix $\mathbf{B}$ partitioned as follows:

$$\mathbf{B} = \left[\begin{array}{c|c} \underset{N \times N}{\mathbf{0}} & \underset{N \times M}{\mathbf{G}} \\ \hline \underset{M \times N}{\mathbf{G}^{\mathrm{T}}} & \underset{M \times M}{\mathbf{0}} \end{array}\right] \begin{array}{c} \Uparrow \\ N \\ \Downarrow \\ \Uparrow \\ M \\ \Downarrow \end{array} \qquad (6.2)$$

$$|\Leftarrow N \Rightarrow||\Leftarrow M \Rightarrow|$$

B is *Hermitian* because

$$\mathbf{B}^{\mathrm{T}} = \mathbf{B} \qquad (6.3)$$

Note, for example,

$$B_{1,N+3} = G_{13} \qquad (6.4)$$

and

$$B_{N+3,1} = (\mathbf{G}^{\mathrm{T}})_{31} = G_{13} \text{ , etc.} \qquad (6.5)$$

6.2.2 The Role of $\mathbf{G}^{\mathrm{T}}$ as an Operator

Analogous to Equation (1.13), we can define an equation for $\mathbf{G}^{\mathrm{T}}$ as follows:

$$\underset{M \times N}{\mathbf{G}^{\mathrm{T}}} \quad \underset{N \times 1}{\mathbf{y}} = \underset{M \times 1}{\mathbf{c}} \qquad (6.6)$$

We do not have to have a particular **y** and **c** in mind when we do this. We are simply interested in the mapping of an *N*-dimensional vector into an *M*-dimensional vector by $\mathbf{G}^{\mathrm{T}}$.

We can combine $\mathbf{Gm} = \mathbf{d}$ and $\mathbf{G}^{\mathrm{T}}\mathbf{y} = \mathbf{c}$, using **B**, as

$$\left[\begin{array}{c|c} \mathbf{0} & \mathbf{G} \\ \hline \mathbf{G}^{\mathrm{T}} & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \mathbf{y} \\ \hline \mathbf{m} \end{array}\right] = \left[\begin{array}{c} \mathbf{d} \\ \hline \mathbf{c} \end{array}\right] \qquad (6.7)$$

or

$$\underset{(N+M) \times (N+M)}{\mathbf{B}} \quad \underset{(N+M) \times 1}{\mathbf{z}} = \underset{(N+M) \times 1}{\mathbf{b}} \qquad (6.8)$$

where we have

$$\mathbf{z} = \left[\begin{array}{c} \mathbf{y} \\ \hline \mathbf{m} \end{array} \right] \tag{6.9}$$

and

$$\mathbf{b} = \left[\begin{array}{c} \mathbf{d} \\ \hline \mathbf{c} \end{array} \right] \tag{6.10}$$

Note that $\mathbf{z}$ and $\mathbf{b}$ are both $(N + M) \times 1$ column vectors.

6.3 The Eigenvalue Problem for B

The eigenvalue problem for the $(N + M) \times (N + M)$ matrix $\mathbf{B}$ is given by

$$\mathbf{B}\mathbf{w}_i = \eta_i \mathbf{w}_i \qquad i = 1, 2, \ldots, N + M \tag{6.11}$$

6.3.1 Properties of B

The matrix $\mathbf{B}$ is Hermitian. Therefore, all $N + M$ eigenvalues η_i are real. In preparation for solving the eigenvalue problem, we define the eigenvector matrix $\mathbf{W}$ for $\mathbf{B}$ as follows:

$$\underset{(N+M)\times(N+M)}{\mathbf{W}} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{w}_1 & \mathbf{w}_2 & \cdots & \mathbf{w}_{N+M} \\ \vdots & \vdots & & \vdots \end{bmatrix} \tag{6.12}$$

We note that $\mathbf{W}$ is an orthogonal matrix, and thus

$$\mathbf{W}^{\mathrm{T}}\mathbf{W} = \mathbf{W}\mathbf{W}^{\mathrm{T}} = \mathbf{I}_{N+M} \tag{6.13}$$

This is equivalent to

$$\mathbf{w}_i^{\mathrm{T}} \mathbf{w}_j = \delta_{ij} \tag{6.14}$$

where $\mathbf{w}_i$ is the ith eigenvector in $\mathbf{W}$.

6.3.2 Partitioning W

Each eigenvector $\mathbf{w}_i$ is $(N + M) \times 1$. Consider partitioning $\mathbf{w}_i$ such that

$$\mathbf{w}_i = \begin{bmatrix} \mathbf{u}_i \\ \hline \mathbf{v}_i \end{bmatrix} \begin{matrix} \Uparrow \\ N \\ \Downarrow \\ \Uparrow \\ M \\ \Downarrow \end{matrix} \tag{6.15}$$

That is, we "stack" an N-dimensional vector $\mathbf{u}_i$ and an M-dimensional vector $\mathbf{v}_i$ into a single $(N + M)$-dimensional vector.

Then the eigenvalue problem for $\mathbf{B}$ from Equation (6.11) becomes

$$\left[\begin{array}{c|c} \mathbf{0} & \mathbf{G} \\ \hline \mathbf{G}^{\mathrm{T}} & \mathbf{0} \end{array}\right] \begin{bmatrix} \mathbf{u}_i \\ \hline \mathbf{v}_i \end{bmatrix} = \eta_i \begin{bmatrix} \mathbf{u}_i \\ \hline \mathbf{v}_i \end{bmatrix} \tag{6.16}$$

This can be written as

$$\underset{N \times M}{\mathbf{G}} \; \underset{M \times 1}{\mathbf{v}_i} = \underset{N \times 1}{\eta_i \mathbf{u}_i} \qquad i = 1, 2, \ldots, N + M \tag{6.17}$$

and

$$\underset{M \times N}{\mathbf{G}^{\mathrm{T}}} \; \underset{N \times 1}{\mathbf{u}_i} = \underset{M \times 1}{\eta_i \mathbf{v}_i} \qquad i = 1, 2, \ldots, N + M \tag{6.18}$$

Equations (6.17) and (6.18) together are called the *shifted eigenvalue problem for* **G**. It is not an eigenvalue problem for **G**, since **G** is not square and eigenvalue problems are only defined for square matrices. Still, it is analogous to an eigenvalue problem. Note that **G** operates on an M-dimensional vector and returns an N-dimensional vector. $\mathbf{G}^{\mathrm{T}}$ operates on an N-dimensional vector and returns an M-dimensional vector. Furthermore, the vectors are shared by **G** and $\mathbf{G}^{\mathrm{T}}$.

6.4 Solving the Shifted Eigenvalue Problem

Equations (6.17) and (6.18) can be solved by combining them into two related eigenvalue problems involving $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ and $\mathbf{G}\mathbf{G}^{\mathrm{T}}$, respectively.

6.4.1 The Eigenvalue Problem for $\mathbf{G}^{\mathrm{T}}\mathbf{G}$

Eigenvalue problems are only defined for square matrices. Note, then, that $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ is $M \times M$, and hence has an eigenvalue problem. The procedure is as follows:

Starting with Equation (6.18)

$$\mathbf{G}^{\mathrm{T}}\mathbf{u}_i = \eta_i \mathbf{v}_i \tag{6.18}$$

Multiply both sides by η_i

$$\eta_i \mathbf{G}^{\mathrm{T}}\mathbf{u}_i = \eta_i^2 \mathbf{v}_i \tag{6.19}$$

or

$$\mathbf{G}^{\mathrm{T}}(\eta_i \mathbf{u}_i) = \eta_i^2 \mathbf{v}_i \tag{6.20}$$

But, by Equation (6.17), we have

$$\eta_i \mathbf{u}_i = \mathbf{G}\mathbf{v}_i \tag{6.17}$$

Thus

$$\boxed{\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{v}_i = \eta_i^2 \mathbf{v}_i \qquad i = 1, 2, \ldots, M} \tag{6.21}$$

This is just the eigenvalue problem for $\mathbf{G}^{\mathrm{T}}\mathbf{G}$! We were able to manipulate the shifted eigenvalue problem into an eigenvalue problem that, presumably, we can solve.

We make the following notes:

1. $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ is Hermitian.

$$(\mathbf{G}^{\mathrm{T}}\mathbf{G})_{ij} = \sum_{k=1}^{N} (G^{\mathrm{T}})_{ik} G_{kj} = \sum_{k=1}^{N} g_{ki} g_{kj} \tag{6.22}$$

$$(\mathbf{G}^{\mathrm{T}}\mathbf{G})_{ji} = \sum_{k=1}^{N} (G^{\mathrm{T}})_{jk} G_{ki} = \sum_{k=1}^{N} g_{kj} g_{ki} = (\mathbf{G}^{\mathrm{T}}\mathbf{G})_{ij} \tag{6.23}$$

2. Therefore, all M η_i^2 are real. Because the diagonal entries of $\mathbf{G}^T\mathbf{G}$ are all ≥ 0, then all η_i^2 are also ≥ 0. This means that $\mathbf{G}^T\mathbf{G}$ is positive semidefinite (one definition of which is, simply, that all the eigenvalues are real and ≥ 0).

We can combine the M equations implied by Equation (6.20) into matrix notation as

$$\underset{M \times M}{\mathbf{G}^T\mathbf{G}} \underset{M \times M}{\mathbf{V}} = \underset{M \times M}{\mathbf{V}} \underset{M \times M}{\mathbf{M}} \tag{6.24}$$

where $\mathbf{V}$ is defined as follows:

$$\mathbf{V} = \underset{M \times M}{\begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_M \\ \vdots & \vdots & & \vdots \end{bmatrix}} \tag{6.25}$$

and

$$\mathbf{M} = \underset{M \times M}{\begin{bmatrix} \eta_1^2 & 0 & \cdots & 0 \\ 0 & \eta_2^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \eta_M^2 \end{bmatrix}} \tag{6.26}$$

3. Because $\mathbf{G}^T\mathbf{G}$ is a Hermitian matrix, $\mathbf{V}$ is itself an orthogonal matrix:

$$\mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}_M \tag{6.27}$$

6.4.2 The Eigenvalue Problem for $\mathbf{GG}^T$

The procedure for forming the eigenvalue problem for $\mathbf{GG}^T$ is very analogous to that of $\mathbf{G}^T\mathbf{G}$. We note that $\mathbf{GG}^T$ is $N \times N$. Starting with Equation (6.17),

$$\mathbf{G}\mathbf{v}_i = \eta_i \mathbf{u}_i \tag{6.17}$$

Again, multiply by η_i

$$\mathbf{G}(\eta_i \mathbf{v}_i) = \eta_i^2 u_i \tag{6.28}$$

But by Equation (6.18), we have

$$\mathbf{G}^T\mathbf{u}_i = \eta_i \mathbf{v}_i \tag{6.18}$$

Thus

$$\boxed{\mathbf{G}\mathbf{G}^{\mathrm{T}}\mathbf{u}_i = \eta_i^2 \mathbf{u}_i \qquad i = 1, 2, \ldots, N} \tag{6.29}$$

We make the following notes for this eigenvalue problem:

1. $\mathbf{GG}^{\mathrm{T}}$ is Hermitian.

2. $\mathbf{GG}^{\mathrm{T}}$ is positive semidefinite.

3. Combining the N equations in Equation (6.29), we have

$$\mathbf{GG}^{\mathrm{T}}\mathbf{U} = \mathbf{UN} \tag{6.30}$$

where

$$\mathbf{U} = \underset{N \times N}{\begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_N \\ \vdots & \vdots & & \vdots \end{bmatrix}} \tag{6.31}$$

and

$$\mathbf{N} = \underset{N \times N}{\begin{bmatrix} \eta_1^2 & 0 & \cdots & 0 \\ 0 & \eta_2^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \eta_N^2 \end{bmatrix}} \tag{6.32}$$

4. $\mathbf{U}$ is an orthogonal matrix

$$\mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{UU}^{\mathrm{T}} = \mathbf{I}_N \tag{6.33}$$

6.5 How Many η_i Are There, Anyway??

A careful look at Equations (6.11), (6.21), and (6.29) shows that the eigenvalue problems for $\mathbf{B}$, $\mathbf{G}^{\mathrm{T}}\mathbf{G}$, and $\mathbf{GG}^{\mathrm{T}}$ are defined for $(N + M)$, M, and N values of i, respectively. Just how many η_i are there?

6.5.1 Introducing *P*, the Number of Nonzero Pairs ($+\eta_i, -\eta_i$)

Equation (6.11)

$$\mathbf{B}\mathbf{w}_i = \eta_i \mathbf{w}_i \tag{6.11}$$

can be used to determine $(N + M)$ real η_i. Equation (6.21),

$$\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{v}_i = \eta_i^2 \mathbf{v}_i \tag{6.21}$$

can be used to determine M real η_i^2 since $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ is $M \times M$. Equation (6.29)

$$\mathbf{G}\mathbf{G}^{\mathrm{T}}\mathbf{u}_i = \eta_i^2 \mathbf{u}_i \tag{6.29}$$

can be used to determine N real η_i^2 since $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ is $N \times N$.

This section will convince you, I hope, that the following are true:

1. There are P pairs of nonzero η_i, where each pair consists of $(+\eta_i, -\eta_i)$.

2. If $+\eta_i$ is an eigenvalue of

$$\mathbf{B}\mathbf{w}_i = \eta_i \mathbf{w}_i \tag{6.11}$$

   and the associated eigenvector $\mathbf{w}_i$ is given by

$$\mathbf{w}_i = \begin{bmatrix} \mathbf{u}_i \\ \mathbf{v}_i \end{bmatrix} \tag{6.34}$$

   then the eigenvector associated with $-\eta_i$ is given by

$$\mathbf{w}_i' = \begin{bmatrix} -\mathbf{u}_i \\ \mathbf{v}_i \end{bmatrix} \tag{6.35}$$

3. There are $(N + M) - 2P$ zero η_i.

4. You can know everything you need to know about the shifted eigenvalue problem by retaining *only* the information associated with the positive η_i.

5. P is less than or equal to the minimum of N and M.

$$P \le \min(N, M) \tag{6.36}$$

6.5.2 Finding the Eigenvector Associated With $-\eta_i$

Suppose that you have found $\mathbf{w}_i$, a solution to Equation (6.11) associated with η_i. It also satisfies the shifted eigenvalue problem

$$\mathbf{G}\mathbf{v}_i = \eta_i \mathbf{u}_i \tag{6.17}$$

and

$$\mathbf{G}^{\mathrm{T}}\mathbf{u}_i = \eta_i \mathbf{v}_i \tag{6.18}$$

Let us try $-\eta_i$ as an eigenvalue and $\mathbf{w}'_i$ given by

$$\mathbf{w}'_i = \begin{bmatrix} -\mathbf{u}_i \\ \mathbf{v}_i \end{bmatrix} \tag{6.35}$$

and see if it satisfies Equations (6.16) and (6.17)

$$\mathbf{G}\mathbf{v}_i = (-\eta_i)(-\mathbf{u}_i) = \eta_i \mathbf{u}_i \tag{6.37}$$

and

$$\mathbf{G}^{\mathrm{T}}(-\mathbf{u}_i) = (-\eta_i)\mathbf{v}_i \tag{6.38}$$

or

$$\mathbf{G}^{\mathrm{T}}\mathbf{u}_i = \eta_i \mathbf{v}_i \tag{6.39}$$

From this we conclude that the nonzero eigenvalues of **B** come in pairs. The relationship between the solutions is given in Equations (6.34) and (6.35). *Note that this property of paired eigenvalues and eigenvectors is not the case for the general eigenvalue problem.* It results from the symmetry of the shifted eigenvalue problem.

6.5.3 No New Information From the $-\eta_i$ System

Let us form an ordered eigenvalue matrix **D** for **B** given by (next page)

$$\mathbf{D} = \begin{bmatrix} \eta_1 & 0 & & & & \cdots & & & & 0 \\ 0 & \eta_2 & & & & & & & & \\ & & \ddots & & & & & & & \\ & & & \eta_P & & & & & & \\ & & & & -\eta_1 & & & & & \\ \vdots & & & & & -\eta_2 & & & & \vdots \\ & & & & & & \ddots & & & \\ & & & & & & & -\eta_P & & \\ & & & & & & & & 0 & \\ & & & & & & & & & \ddots \\ 0 & & & & & \cdots & & & & 0 \end{bmatrix}_{(N+M)\times(N+M)} \tag{6.40}$$

where $\eta_1 \geq \eta_2 \geq \cdots \geq \eta_P$. Note that the ordering of matrices in eigenvalue problems is arbitrary, but must be internally consistent. Then the eigenvalue problem for **B** from Equation (6.11) becomes

$$\mathbf{BW} = \mathbf{WD} \tag{6.41}$$

where now the $(N + M) \times (N + M)$ dimensional matrix **W** is given by

$$\mathbf{W} = \left[\begin{array}{cccccccccc} \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_P & -\mathbf{u}_1 & -\mathbf{u}_2 & \cdots & -\mathbf{u}_P & \mathbf{u}_{2P+1} & \cdots & \mathbf{u}_{N+M} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \hline \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_P & \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_P & \mathbf{v}_{2P+1} & \cdots & \mathbf{v}_{N+M} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \end{array}\right] \tag{6.42}$$

$$|\Leftarrow \quad P \quad \Rightarrow|\,|\Leftarrow \quad P \quad \Rightarrow|\,|\Leftarrow (N+M) - 2P \Rightarrow|$$

The second P eigenvectors certainly contain independent information about the eigenvectors $\mathbf{w}_i$ in $(N + M)$-space. They contain *no new* information, however, about $\mathbf{u}_i$ or $\mathbf{v}_i$, in N- and M-space, respectively, since $-\mathbf{u}_i$ contains no information not already contained in $+\mathbf{u}_i$.

6.5.4 What About the Zero Eigenvalues η_i's, $i = (2P + 1), \ldots, N + M$?

For the zero eigenvalues, the shifted eigenvalue problem becomes

$$\mathbf{Gv}_i = \eta_i \mathbf{u}_i = 0\mathbf{u}_i = \underset{N \times 1}{\mathbf{0}} \qquad i = (2P + 1), \ldots, (N + M) \tag{6.43}$$

and

$$\mathbf{G}^{\mathrm{T}}\mathbf{u}_i = \eta_i \mathbf{v}_i = 0\mathbf{v}_i = \underset{M \times 1}{\mathbf{0}} \qquad i = (2P+1), \ldots, (N+M) \tag{6.44}$$

where $\mathbf{0}$ is a vector of zeros of the appropriate dimension.

If you premultiply Equation (6.43) by $\mathbf{G}^{\mathrm{T}}$ and Equation (6.44) by $\mathbf{G}$, you obtain

$$\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{v}_i = \mathbf{G}^{\mathrm{T}}\mathbf{0} = \underset{(M \times 1)}{\mathbf{0}} \tag{6.45}$$

and

$$\mathbf{G}\mathbf{G}^{\mathrm{T}}\mathbf{u}_i = \mathbf{G}\mathbf{0} = \underset{(N \times 1)}{\mathbf{0}} \tag{6.46}$$

Therefore, we conclude that the $\mathbf{u}_i$, $\mathbf{v}_i$ associated with zero η_i for $\mathbf{B}$ are simply the eigenvectors of $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ and $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ associated with zero eigenvalues for $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ and $\mathbf{G}^{\mathrm{T}}\mathbf{G}$, respectively!

6.5.5 How Big is P?

Now that we have seen that the eigenvalues come in P pairs of nonzero values, how can we determine the size of P? We will see that you can determine P from either $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ or $\mathbf{G}\mathbf{G}^{\mathrm{T}}$, and that P is bounded by the smaller of N and M, the number of observations and model parameters, respectively. The steps are as follows.

Step 1. Let the number of nonzero eigenvalues η_i^2 of $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ be P. Since $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ is $M \times M$, there are only M η_i^2 all together. Thus, P is less than or equal to M.

Step 2. If $\eta_i^2 \neq 0$ is an eigenvalue of $\mathbf{G}^{\mathrm{T}}\mathbf{G}$, then it is also an eigenvalue of $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ since

$$\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{v}_i = \eta_i^2 \mathbf{v}_i \tag{6.21}$$

and

$$\mathbf{G}\mathbf{G}^{\mathrm{T}}\mathbf{u}_i = \eta_i^2 \mathbf{u}_i \tag{6.29}$$

Thus the nonzero η_i's are shared by $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ and $\mathbf{G}\mathbf{G}^{\mathrm{T}}$.

Step 3. P is less than or equal to N since $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ is $N \times N$. Therefore, since $P \leq M$ and $P \leq N$,

$$P \leq \min(N, M)$$

Thus, to determine P, you can do the eigenvalue problem for either $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ $(M \times M)$ or $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ $(N \times N)$. It makes sense to choose the smaller of the two matrices. That is, one chooses $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ if $M < N$, or $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ if $N < M$.

6.6 Introducing Singular Values

6.6.1 Introduction

Recalling Equation (6.30) defining the eigenvalue problem for $\mathbf{GG}^T$

$$\underset{N\times N}{\mathbf{GG}^T}\ \underset{N\times N}{\mathbf{U}} = \underset{N\times N}{\mathbf{U}}\ \underset{N\times N}{\mathbf{N}} \tag{6.30}$$

The matrix $\mathbf{U}$ contains the eigenvectors $\mathbf{u}_i$, and can be ordered as

$$\underset{N\times N}{\mathbf{U}} = \left[\begin{array}{cccc|ccc} \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_P & \mathbf{u}_{P+1} & \cdots & \mathbf{u}_N \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \end{array}\right] \begin{array}{c} \Uparrow \\ N \\ \Downarrow \end{array} \tag{6.47}$$
$$|\Leftarrow \quad P \quad \Rightarrow|\,|\Leftarrow \ N-P \ \Rightarrow|$$

or

$$\mathbf{U} = [\,\mathbf{U}_P \mid \mathbf{U}_0\,] \tag{6.48}$$

Recall Equation (6.24), which defined the eigenvalue problem for $\mathbf{G}^T\mathbf{G}$,

$$\underset{M\times M}{\mathbf{G}^T\mathbf{G}}\ \underset{M\times M}{\mathbf{V}} = \underset{M\times M}{\mathbf{V}}\ \underset{M\times M}{\mathbf{M}} \tag{6.24}$$

The matrix $\mathbf{V}$ of eigenvectors is given by

$$\underset{M\times M}{\mathbf{V}} = \left[\begin{array}{cccc|ccc} \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_P & \mathbf{v}_{P+1} & \cdots & \mathbf{v}_M \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \end{array}\right] \begin{array}{c} \Uparrow \\ M \\ \Downarrow \end{array} \tag{6.49}$$
$$|\Leftarrow \quad P \quad \Rightarrow|\,|\Leftarrow \ M-P \ \Rightarrow|$$

or

$$\mathbf{V} = [\,\mathbf{V}_P \mid \mathbf{V}_0\,] \tag{6.50}$$

where the $\mathbf{u}_i$, $\mathbf{v}_i$ satisfy

$$\mathbf{Gv}_i = \eta_i \mathbf{u}_i \tag{6.17}$$

and

$$\mathbf{G}^T\mathbf{u}_i = \eta_i \mathbf{v}_i \tag{6.18}$$

and where we have chosen the P positive η_i from

$$\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{v}_i = \eta_i^2\mathbf{v}_i \tag{6.21}$$

$$\mathbf{G}\mathbf{G}^{\mathrm{T}}\mathbf{u}_i = \eta_i^2\mathbf{u}_i \tag{6.29}$$

Note that it is customary to order the $\mathbf{u}_i$, $\mathbf{v}_i$ such that

$$\eta_1 \geq \eta_2 \geq \cdots \geq \eta_P \tag{6.51}$$

6.6.2 Definition of the Singular Value

We define a singular value λ_i from Equation (6.21) or (6.29) as the positive square root of the eigenvalue η_i^2 for $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ or $\mathbf{G}\mathbf{G}^{\mathrm{T}}$. That is,

$$\lambda_i = +\sqrt{\eta_i^2} \tag{6.52}$$

Singular values are not eigenvalues. λ_i is not an eigenvalue for $\mathbf{G}$ or $\mathbf{G}^{\mathrm{T}}$, since the eigenvalue problem is not defined for $\mathbf{G}$ or $\mathbf{G}^{\mathrm{T}}$, $N \neq M$. They are, of course, eigenvalues for $\mathbf{B}$ in Equation (6.11), but we will never explicitly deal with $\mathbf{B}$. The matrix $\mathbf{B}$ is a construct that allowed us to formulate the shifted eigenvalue problem, but in practice, it is never formed. Nevertheless, you will often read, or hear, λ_i referred to as an eigenvalue.

6.6.3 Definition of Λ, the Singular-Value Matrix

We can form an $N \times M$ matrix with the singular values on the diagonal. If $M > N$, it has the form

$$\underset{N\times M}{\Lambda} = \left[\begin{array}{cccccc|c} \lambda_1 & 0 & & \cdots & & 0 & \\ 0 & \lambda_2 & & & & & \\ & & \ddots & & & & \\ \vdots & & & \lambda_P & & \vdots & \mathbf{0} \\ & & & & 0 & & \\ & & & & & \ddots & \\ 0 & & & \cdots & & 0 & \end{array}\right] \begin{array}{c}\Uparrow \\ \\ \\ \mathrm{N} \\ \\ \\ \Downarrow\end{array} \tag{6.53}$$
$$|\Leftarrow \quad N \quad \Rightarrow||\Leftarrow M-N \Rightarrow|$$

If $N > M$, it has the form

$$\underset{N\times M}{\Lambda} = \left[\begin{array}{cccccc} \lambda_1 & 0 & & \cdots & & 0 \\ 0 & \lambda_2 & & & & \\ & & \ddots & & & \\ \vdots & & & \lambda_P & & \vdots \\ & & & & 0 & \\ & & & & & \ddots \\ 0 & & & \cdots & & 0 \\ \hline & & & & & \\ & & & \mathbf{0} & & \\ & & & & & \end{array}\right] \begin{array}{c} \overline{\Uparrow} \\ \\ M \\ \\ \underline{\Downarrow} \\ \overline{\Uparrow} \\ N-M \\ \underline{\Downarrow} \end{array} \tag{6.54}$$

$$|\!\Leftarrow \qquad M \qquad \Rightarrow\!|$$

Then the shifted eigenvalue problem

$$\mathbf{G}\mathbf{v}_i = \eta_i \mathbf{u}_i \tag{6.17}$$

and

$$\mathbf{G}^{\mathrm{T}}\mathbf{u}_i = \eta_i \mathbf{v}_i \tag{6.18}$$

can be written as

$$\mathbf{G}\mathbf{v}_i = \lambda_i \mathbf{u}_i \tag{6.55}$$

and

$$\mathbf{G}^{\mathrm{T}}\mathbf{u}_i = \lambda_i \mathbf{v}_i \tag{6.56}$$

where η_i has been replaced by λ_i since all information about **U**, **V** can be obtained from the positive η_i.

Equations (6.55) and (6.56) can be written in matrix notation as

$$\boxed{\underset{N\times M}{\mathbf{G}} \quad \underset{M\times M}{\mathbf{V}} = \underset{N\times N}{\mathbf{U}} \quad \underset{N\times M}{\Lambda}} \tag{6.57}$$

and

$$\boxed{\underset{M\times N}{\mathbf{G}^{\mathrm{T}}} \quad \underset{N\times N}{\mathbf{U}} = \underset{M\times M}{\mathbf{V}} \quad \underset{M\times N}{\Lambda^{\mathrm{T}}}} \tag{6.58}$$

6.7 Derivation of the Fundamental Decomposition Theorem for General G ($N \times M, N \neq M$)

Recall that we used the eigenvalue problem for square **A** and $\mathbf{A}^T$ to derive a decomposition theorem for square matrices:

$$\mathbf{A} = \mathbf{S}\Lambda\mathbf{R}^T \tag{5.90}$$

where **S**, **R**, and Λ are eigenvector and eigenvalue matrices associated with **A** and $\mathbf{A}^T$. We are now ready to derive an analogous decomposition theorem for the general $N \times M$, $N \neq M$ matrix **G**.

We start with Equation (6.57)

$$\underset{N\times M}{\mathbf{G}} \; \underset{M\times M}{\mathbf{V}} = \underset{N\times N}{\mathbf{U}} \; \underset{N\times M}{\Lambda} \tag{6.57}$$

postmultiply by $\mathbf{V}^T$

$$\mathbf{G}\mathbf{V}\mathbf{V}^T = \mathbf{U}\Lambda\mathbf{V}^T \tag{6.59}$$

But **V** is an orthogonal matrix. That is,

$$\mathbf{V}\mathbf{V}^T = \mathbf{I}_M \tag{6.27}$$

since $\mathbf{G}^T\mathbf{G}$ is Hermitian, and the eigenvector matrices of Hermitian matrices are orthogonal. Therefore, we have the fundamental decomposition theorem for a general matrix **G** given by

$$\boxed{\underset{N\times M}{\mathbf{G}} = \underset{N\times N}{\mathbf{U}} \; \underset{N\times M}{\Lambda} \; \underset{M\times M}{\mathbf{V}^T}} \tag{6.60}$$

By taking the transpose of Equation (6.60), we obtain also

$$\boxed{\underset{M\times N}{\mathbf{G}^T} = \underset{M\times M}{\mathbf{V}} \; \underset{M\times N}{\Lambda^T} \; \underset{N\times N}{\mathbf{U}^T}} \tag{6.61}$$

where

$$\underset{N\times N}{\mathbf{U}} = \begin{bmatrix} \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{u}_1 & \cdots & \mathbf{u}_P & \mathbf{u}_{P+1} & \cdots & \mathbf{u}_N \\ \vdots & & \vdots & \vdots & & \vdots \end{bmatrix} = \left[\, \mathbf{U}_P \;\middle|\; \mathbf{U}_0 \,\right] \tag{6.62}$$

and

$$\underset{M\times M}{\mathbf{V}} = \begin{bmatrix} \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{v}_1 & \cdots & \mathbf{v}_P & \mathbf{v}_{P+1} & \cdots & \mathbf{v}_M \\ \vdots & & \vdots & \vdots & & \vdots \end{bmatrix} = [\ \mathbf{V}_P \ \ | \ \ \mathbf{V}_0\ ] \tag{6.63}$$

and

$$\underset{N\times M}{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & & \cdots & & 0 \\ 0 & \lambda_2 & & & & \\ & & \ddots & & & \\ \vdots & & & \lambda_P & & \vdots \\ & & & & 0 & \\ & & & & & \ddots \\ 0 & & & \cdots & & 0 \end{bmatrix} \tag{6.64}$$

6.8 Singular-Value Decomposition (SVD)

6.8.1 Derivation of Singular-Value Decomposition

We will see below that **G** can be decomposed without any knowledge of the parts of **U** or **V** associated with zero singular values λ_i, i > P. We start with the fundamental decomposition theorem

$$\mathbf{G} = \mathbf{U}\Lambda\mathbf{V}^{\mathrm{T}} \tag{6.60}$$

Let us introduce a $P \times P$ singular-value matrix Λ_P that is a subset of Λ:

$$\Lambda_P = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_P \end{bmatrix} \tag{6.65}$$

We now write out Equation (6.60) in terms of the partitioned matrices as

$$\mathbf{G} = \underset{\Downarrow}{\overset{\Uparrow}{N}} \left[\begin{array}{c|c} \mathbf{U}_P & \mathbf{U}_0 \end{array} \right] \begin{array}{c} \Uparrow \\ P \\ \Downarrow \\ \Uparrow \\ N-P \\ \Downarrow \end{array} \left[\begin{array}{c|c} \Lambda_P & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right] \left[\begin{array}{c} \mathbf{V}_P^{\mathrm{T}} \\ \hline \mathbf{V}_0^{\mathrm{T}} \end{array} \right] \begin{array}{c} \Uparrow \\ P \\ \Downarrow \\ \Uparrow \\ M-P \\ \Downarrow \end{array} \tag{6.66}$$

$$|\Leftarrow \; P \; \Rightarrow| \, |\Leftarrow N\text{–}P \Rightarrow| \qquad |\Leftarrow \; P \; \Rightarrow||\Leftarrow M-P \Rightarrow||\Leftarrow \; M \; \Rightarrow|$$

$$= \underset{\Downarrow}{\overset{\Uparrow}{N}} \left[\begin{array}{c|c} \mathbf{U}_P \Lambda_P & \mathbf{0} \end{array} \right] \begin{bmatrix} \mathbf{V}_P^{\mathrm{T}} \\ \mathbf{V}_0^{\mathrm{T}} \end{bmatrix} \tag{6.67}$$

$$|\Leftarrow \; P \; \Rightarrow||\Leftarrow M\text{–}P \Rightarrow|$$

$$= \mathbf{U}_P \Lambda_P \mathbf{V}_P^{\mathrm{T}} \tag{6.68}$$

That is, we can write **G** as

$$\boxed{\begin{array}{cccc} \mathbf{G} & = \mathbf{U}_P & \Lambda_P & \mathbf{V}_P^{\mathrm{T}} \\ N \times M & N \times P & P \times P & P \times M \end{array}} \tag{6.69}$$

Equation (6.69) is known as the *Singular-Value Decomposition Theorem for* **G**.

The matrices in Equation (6.69) are

1. **G** = an arbitrary $N \times M$ matrix.

2. The eigenvector matrix $\mathbf{U}_P$

$$\mathbf{U}_P = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_P \\ \vdots & \vdots & & \vdots \end{bmatrix} \tag{6.70}$$

where $\mathbf{u}_i$ are the P N-dimensional eigenvectors of

$$\mathbf{G}\mathbf{G}^{\mathrm{T}}\mathbf{u}_i = \eta_i^2 \, \mathbf{u}_i \tag{6.29}$$

associated with nonzero singular values λ_i.

3. The eigenvector matrix $\mathbf{V}_P$

$$\mathbf{V}_P = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_P \\ \vdots & \vdots & & \vdots \end{bmatrix} \tag{6.71}$$

where $\mathbf{v}_i$ are the P M-dimensional eigenvectors of

$$\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{v}_i = \eta_i^2 \mathbf{v}_{\mathrm{i}} \tag{6.21}$$

associated with nonzero singular values λ_i, and

4. The singular-value matrix Λ

$$\Lambda_P = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_P \end{bmatrix} \tag{6.65}$$

where λ_i is the nonzero singular value associated with $\mathbf{u}_i$ and $\mathbf{v}_i$, $i = 1, \ldots, P$.

6.8.2 Rewriting the Shifted Eigenvalue Problem

Now that we have seen that we can reconstruct **G** using only the subsets of **U, V**, and Λ defined in Equations (6.65), (6.69), and (6.71), we can rewrite the shifted eigenvalue problem given by Equations (6.57) and (6.58):

$$\underset{N \times M}{\mathbf{G}} \; \underset{M \times M}{\mathbf{V}} = \underset{N \times N}{\mathbf{U}} \; \underset{N \times M}{\Lambda} \tag{6.57}$$

and

$$\underset{M \times N}{\mathbf{G}^{\mathrm{T}}} \; \underset{N \times N}{\mathbf{U}} = \underset{M \times M}{\mathbf{V}} \; \underset{M \times N}{\Lambda^{\mathrm{T}}} \tag{6.58}$$

as

1. $$\underset{N \times M}{\mathbf{G}} \; \underset{M \times P}{\mathbf{V}_P} = \underset{N \times P}{\mathbf{U}_P} \; \underset{P \times P}{\Lambda_P} \tag{6.72}$$

2. $$\underset{M \times N}{\mathbf{G}^{\mathrm{T}}} \; \underset{N \times P}{\mathbf{U}_{\mathrm{P}}} = \underset{M \times P}{\mathbf{V}_{\mathrm{P}}} \; \underset{P \times P}{\Lambda_{\mathrm{P}}} \tag{6.73}$$

3. $$\underset{N \times M}{\mathbf{G}} \; \underset{M \times (M-P)}{\mathbf{V}_0} = \underset{N \times (M-P)}{\mathbf{0}} \tag{6.74}$$

4. $$\underset{M \times N}{\mathbf{G}^T} \quad \underset{N \times (N-P)}{\mathbf{U}_0} \quad = \quad \underset{M \times (N-P)}{\mathbf{0}} \tag{6.75}$$

Note that the eigenvectors in **V** are a set of *M* orthogonal vectors which span *model space*, while the eigenvectors in **U** are a set of *N* orthogonal vectors which span *data space*. The *P* vectors in $\mathbf{V}_P$ span a *P*-dimensional subset of model space, while the *P* vectors in $\mathbf{U}_P$ span a *P*-dimensional subset of data space. $\mathbf{V}_0$ and $\mathbf{U}_0$ are called *null*, or *zero*, spaces. They are $(M - P)$ and $(N - P)$ dimensional subsets of model and data spaces, respectively.

6.8.3 Summarizing SVD

In summary, we started with Equations (1.13) and (6.5)

$$\mathbf{Gm} = \mathbf{d} \tag{1.13}$$

and

$$\mathbf{G}^T\mathbf{y} = \mathbf{c} \tag{6.6}$$

We constructed

$$\mathbf{B} = \left[\begin{array}{c|c} \underset{N \times N}{\mathbf{0}} & \underset{N \times M}{\mathbf{G}} \\ \hline \underset{M \times N}{\mathbf{G}^T} & \underset{M \times M}{\mathbf{0}} \end{array}\right] \begin{array}{c} \Uparrow \\ N \\ \Downarrow \\ \Uparrow \\ M \\ \Downarrow \end{array} \tag{6.2}$$
$$|\Leftarrow N \Rightarrow||\Leftarrow M \Rightarrow|$$

We then considered the eigenvalue problem for **B**

$$\mathbf{Bw}_i = \eta_i \mathbf{w}_i \qquad i = 1, 2, \ldots, (N + M) \tag{6.11}$$

This led us to the *shifted eigenvalue problem*

$$\mathbf{Gv}_i = \eta_i \mathbf{u}_i \qquad i = 1, 2, \ldots, (N + M) \tag{6.17}$$

and

$$\mathbf{G}^T\mathbf{u}_i = \eta_i \mathbf{v}_i \qquad i = 1, 2, \ldots, (N + M) \tag{6.18}$$

We found that the shifted eigenvalue problem leads us to eigenvalue problems for $\mathbf{G}^T\mathbf{G}$ and $\mathbf{GG}^T$:

$$\mathbf{G}^T\mathbf{G}\mathbf{v}_i = \eta_i^2\mathbf{v}_i \qquad i = 1, 2, \ldots, M \tag{6.21}$$

and

$$\mathbf{G}\mathbf{G}^T\mathbf{u}_i = \eta_i^2\mathbf{u}_i \qquad i = 1, 2, \ldots, N \tag{6.29}$$

We then introduced the singular value λ_i, given by the positive square root of the eigenvalues from Equations (6.20) and (6.28)

$$\lambda_i = +\sqrt{\eta_i^2} \tag{6.52}$$

Equations (6.16), (6.17), (6.20) and (6.28) give us a way to find **U**, **V**, and Λ. They also lead, eventually, to

$$\underset{N \times M}{\mathbf{G}} = \underset{N \times N}{\mathbf{U}} \; \underset{N \times M}{\Lambda} \; \underset{M \times M}{\mathbf{V}^T} \tag{6.60}$$

We then considered partitioning the matrices based on P, the number of nonzero singular values. This led us to *singular-value decomposition*

$$\underset{N \times M}{\mathbf{G}} = \underset{N \times P}{\mathbf{U}_P} \; \underset{P \times P}{\Lambda_P} \; \underset{P \times M}{\mathbf{V}_P^T} \tag{6.76}$$

Before considering an inverse operator based on singular-value decomposition, it is probably useful to cover the mechanics of singular-value decomposition.

6.9 Mechanics of Singular-Value Decomposition

The steps involved in singular-value decomposition are as follows:

Step 1. Begin with $\mathbf{Gm} = \mathbf{d}$.

Form $\mathbf{G}^T\mathbf{G}$ ($M \times M$) or $\mathbf{G}\mathbf{G}^T$ ($N \times N$), whichever is smaller. (N.B. Typically, there are more observations than model parameters; thus, $N > M$, and $\mathbf{G}^T\mathbf{G}$ is the more common choice.)

Step 2. Solve the eigenvalue problem for Hermitian $\mathbf{G}^T\mathbf{G}$ (or $\mathbf{G}\mathbf{G}^T$)

$$\mathbf{G}^T\mathbf{G}\mathbf{v}_i = \eta_i^2\mathbf{v}_i \tag{6.21}$$

1. Find the P nonzero η_i^2 and associated $\mathbf{v}_i$.

2. Let $\lambda_i = +(\eta_i^2)^{1/2}$.

3. Form

$$\mathbf{V}_P = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_P \\ \vdots & \vdots & & \vdots \end{bmatrix} \tag{6.71}$$
$$M \times P$$

and

$$\Lambda_P = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_P \end{bmatrix} \qquad \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_P \tag{6.65}$$
$$P \times P$$

Step 3. Use $\mathbf{G}\mathbf{v}_i = \lambda_i \mathbf{u}_i$ to find $\mathbf{u}_i$ for each known λ_i, $\mathbf{v}_i$.

Note: Finding $\mathbf{u}_i$ this way preserves the *sign* relationship implicit between $\mathbf{u}_i$, $\mathbf{v}_i$ by taking the positive member of each pair $(+\lambda_i, -\lambda_i)$. You will *not* preserve the sign relationship (except by luck) if you use $\mathbf{G}\mathbf{G}^{\mathrm{T}}\mathbf{u}_i = \eta_i^2 \mathbf{u}_i$ to find $\mathbf{u}_i$.

Step 4. Form

$$\mathbf{U}_P = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_P \\ \vdots & \vdots & & \vdots \end{bmatrix} \tag{6.70}$$
$$N \times P$$

Step 5. Finally, form **G** as

$$\underset{N \times M}{\mathbf{G}} = \underset{N \times P}{\mathbf{U}_P} \; \underset{P \times P}{\Lambda_P} \; \underset{P \times M}{\mathbf{V}_P^{\mathrm{T}}} \tag{6.69}$$

6.10 Implications of Singular-Value Decomposition

6.10.1 Relationships Between U, $\mathbf{U}_P$, and $\mathbf{U}_0$

1. $\underset{N \times N}{\mathbf{U}^{\mathrm{T}}} \; \underset{N \times N}{\mathbf{U}} = \underset{N \times N}{\mathbf{U}} \; \underset{N \times N}{\mathbf{U}^{\mathrm{T}}} = \underset{N \times N}{\mathbf{I}_N}$ Since **U** is an orthogonal matrix.

2. $\underset{P \times N}{\mathbf{U}_P^{\mathrm{T}}} \; \underset{N \times P}{\mathbf{U}_P} = \mathbf{I}_P$ $\mathbf{U}_P$ is semiorthogonal because all P vectors in $\mathbf{U}_P$ are perpendicular to each other.

3. $\underset{N \times P}{\mathbf{U}_P} \quad \underset{P \times N}{\mathbf{U}_P^{\mathrm{T}}} \neq \mathbf{I}_N$ — (Unless $P = N$.) $\mathbf{U}_P\mathbf{U}_P^{\mathrm{T}}$ is $N \times N$. $\mathbf{U}_P$ cannot span N-space with only P (N-dimensional) vectors.

4. $\underset{(N-P) \times N}{\mathbf{U}_0^{\mathrm{T}}} \quad \underset{N \times (N-P)}{\mathbf{U}_0} = \mathbf{I}_{N-P}$ — $\mathbf{U}_0$ is semiorthogonal since the $(N - P)$ vectors in $\mathbf{U}_0$ are all perpendicular to each other.

5. $\underset{N \times (N-P)}{\mathbf{U}_0} \quad \underset{(N-P) \times N}{\mathbf{U}_0^{\mathrm{T}}} \neq \mathbf{I}_N$ — $\mathbf{U}_0$ has $(N - P)$ N-dimensional vectors in it. It cannot span N-space. $\mathbf{U}_0\mathbf{U}_0^{\mathrm{T}}$ is $N \times N$.

6. $\underset{P \times N}{\mathbf{U}_P^{\mathrm{T}}} \quad \underset{N \times (N-P)}{\mathbf{U}_0} = \underset{P \times (N-P)}{\mathbf{0}}$ — Since all the eigenvectors in $\mathbf{U}_P$ are perpendicular to all the eigenvectors in $\mathbf{U}_0$.

7. $\underset{(N-P) \times N}{\mathbf{U}_0^{\mathrm{T}}} \quad \underset{N \times P}{\mathbf{U}_{\mathrm{P}}} = \underset{(N-P) \times P}{\mathbf{0}}$ — Again, since all the eigenvectors in $\mathbf{U}_0$ are perpendicular to all the eigenvectors in $\mathbf{U}_P$.

6.10.2 Relationships Between V, $\mathbf{V}_P$, and $\mathbf{V}_0$

1. $\underset{M \times M}{\mathbf{V}^{\mathrm{T}}} \quad \underset{M \times M}{\mathbf{V}} = \underset{M \times M}{\mathbf{V}} \quad \underset{M \times M}{\mathbf{V}^{\mathrm{T}}} = \mathbf{I}_M$ — Since $\mathbf{V}$ is an orthogonal matrix.

2. $\underset{P \times M}{\mathbf{V}_P^{\mathrm{T}}} \quad \underset{M \times P}{\mathbf{V}_P} = \mathbf{I}_P$ — $\mathbf{V}_P$ is semiorthogonal because all P vectors in $\mathbf{V}_P$ are perpendicular to each other.

3. $\underset{M \times P}{\mathbf{V}_P} \quad \underset{P \times M}{\mathbf{V}_P^{\mathrm{T}}} \neq \mathbf{I}_M$ — (Unless $P = M$.) $\mathbf{V}_P^{\mathrm{T}}$ is $M \times M$. $\mathbf{V}_P$ cannot span M-space with only P (M-dimensional) vectors.

4. $\underset{(M-P) \times M}{\mathbf{V}_0^{\mathrm{T}}} \quad \underset{M \times (M-P)}{\mathbf{V}_0} = \mathbf{I}_{M-P}$ — $\mathbf{V}_0$ is semiorthogonal since the $(M - P)$ vectors in $\mathbf{V}_0$ are all perpendicular to each other.

5. $\underset{M \times (M-P)}{\mathbf{V}_0} \quad \underset{(M-P) \times M}{\mathbf{V}_0^{\mathrm{T}}} \neq \mathbf{I}_M$ — $\mathbf{V}_0$ has $(M - P)$ M-dimensional vectors in it. It cannot span M-space. $\mathbf{V}_0\mathbf{V}_0^{\mathrm{T}}$ is $M \times M$.

6. $\underset{P \times M}{\mathbf{V}_P^{\mathrm{T}}} \quad \underset{M \times (M-P)}{\mathbf{V}_0} = \underset{P \times (M-P)}{\mathbf{0}}$ — Since all the eigenvectors in $\mathbf{V}_P$ are perpendicular to all the eigenvectors in $\mathbf{V}_0$.

7. $\underset{(M-P) \times M}{\mathbf{V}_0^{\mathrm{T}}} \quad \underset{M \times P}{\mathbf{V}_P} = \underset{(M-P) \times P}{\mathbf{0}}$ — Again, since all the eigenvectors in $\mathbf{V}_0$ are perpendicular to all the eigenvectors in $\mathbf{V}_P$.

6.10.3 Graphic Representation of U, $\mathbf{U}_P$, $\mathbf{U}_0$, V, $\mathbf{V}_P$, $\mathbf{V}_0$ Spaces

Recall that starting with Equation (1.13)

$$\underset{N \times M}{\mathbf{G}} \quad \underset{M \times 1}{\mathbf{m}} = \underset{N \times 1}{\mathbf{d}} \tag{1.13}$$

we obtained the fundamental decomposition theorem

$$\underset{N \times M}{\mathbf{G}} = \underset{N \times N}{\mathbf{U}} \quad \underset{N \times M}{\Lambda} \quad \underset{M \times M}{\mathbf{V}^{\mathrm{T}}} \tag{6.60}$$

and singular-value decomposition

$$\underset{N \times M}{\mathbf{G}} = \underset{N \times P}{\mathbf{U}_P} \quad \underset{P \times P}{\Lambda_P} \quad \underset{P \times M}{\mathbf{V}_P^{\mathrm{T}}} \tag{6.69}$$

This gives us the following:

1. Recall the definitions of $\mathbf{U}$, $\mathbf{U}_P$, and $\mathbf{U}_0$

$$\underset{N \times N}{\mathbf{U}} = \begin{bmatrix} \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{u}_1 & \cdots & \mathbf{u}_P & \mathbf{u}_{P+1} & \cdots & \mathbf{u}_N \\ \vdots & & \vdots & \vdots & & \vdots \end{bmatrix} = [\ \mathbf{U}_P \ | \ \mathbf{U}_0 \ ] \tag{6.62}$$

2. Similarly, recall the definitions for $\mathbf{V}$, $\mathbf{V}_P$ and $\mathbf{V}_0$

$$\underset{M \times M}{\mathbf{V}} = \begin{bmatrix} \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{v}_1 & \cdots & \mathbf{v}_P & \mathbf{v}_{P+1} & \cdots & \mathbf{v}_N \\ \vdots & & \vdots & \vdots & & \vdots \end{bmatrix} = [\ \mathbf{V}_P \ | \ \mathbf{V}_0 \ ] \tag{6.63}$$

3. Combining $\mathbf{U}$, $\mathbf{U}_P$, $\mathbf{U}_0$, and $\mathbf{V}$, $\mathbf{V}_P$, $\mathbf{V}_0$ graphically

$$\begin{array}{c|c|c|} & |\Leftarrow \ P \ \Rightarrow| & |\Leftarrow \ N-P \ \Rightarrow| \\ \Uparrow \ N \ \Downarrow & \mathbf{U}_P & \mathbf{U}_0 \\ \hline \Uparrow \ M \ \Downarrow & \mathbf{V}_P & \mathbf{V}_0 \\ & |\Leftarrow \ P \ \Rightarrow| & |\Leftarrow \ M-P \ \Rightarrow| \end{array} \tag{6.77}$$

4. Summarizing:

   (1) **V** is an $M \times M$ matrix with the eigenvectors of $\mathbf{G}^T\mathbf{G}$ as columns. It is an orthogonal matrix.

   (2) $\mathbf{V}_P$ is an $M \times P$ matrix with the P eigenvectors of $\mathbf{G}^T\mathbf{G}$ associated with nonzero eigenvalues of $\mathbf{G}^T\mathbf{G}$. $\mathbf{V}_P$ is a semiorthogonal matrix.

   (3) $\mathbf{V}_0$ is an $M \times (M - P)$ matrix with the $M - P$ eigenvectors of $\mathbf{G}^T\mathbf{G}$ associated with the zero eigenvalues of $\mathbf{G}^T\mathbf{G}$. $\mathbf{V}_0$ is a semiorthogonal matrix.

   (4) The eigenvectors in **V**, $\mathbf{V}_P$, or $\mathbf{V}_0$ are all M-dimensional vectors. They are all associated with *model* space, since **m**, the model parameter vector of **Gm = d**, is an M-dimensional vector.

   (5) **U** is an $N \times N$ matrix with the eigenvectors of $\mathbf{G}\mathbf{G}^T$ as columns. It is an orthogonal matrix.

   (6) $\mathbf{U}_P$ is an $N \times P$ matrix with the P eigenvectors of $\mathbf{G}\mathbf{G}^T$ associated with the nonzero eigenvalues of $\mathbf{G}\mathbf{G}^T$. $\mathbf{U}_P$ is a semiorthogonal matrix.

   (7) $\mathbf{U}_0$ is an $N \times (N - P)$ matrix with the $N - P$ eigenvectors of $\mathbf{G}\mathbf{G}^T$ associated with the zero eigenvalues of $\mathbf{G}\mathbf{G}^T$. $\mathbf{U}_0$ is a semiorthogonal matrix.

   (8) The eigenvectors of **U**, $\mathbf{U}_P$ or $\mathbf{U}_0$ are all N-dimensional vectors. They are all associated with *data* space, since **d**, the data vector of **Gm = d**, is an N-dimensional vector.

6.11 Classification of d = Gm Based on *P*, *M*, and *N*

6.11.1 Introduction

In Section 3.3 we introduced a classification of the system of equations

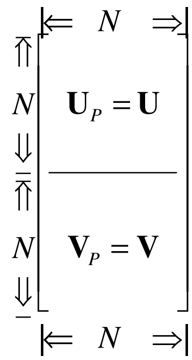
$$\mathbf{d} = \mathbf{Gm} \tag{1.13}$$

based on the dimensions of **d** ($N \times 1$) and **m** ($M \times 1$). I said at the time that I found the classification lacking, and would return to it later. Now that we have considered singular-value decomposition, including finding P, the number of nonzero singular values, I would like to introduce a better classification scheme.

There are four basic classes of problems, based on the relationship between P, M, and N. We will consider each class one at a time below.

6.11.2 Class I: *P* = *M* = *N*

Graphically, for this case, we have (following page):



1. $\mathbf{U}_0$ and $\mathbf{V}_0$ are empty.

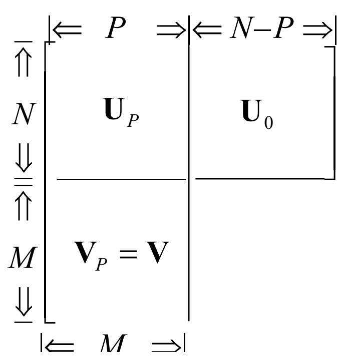
2. **G** has a unique, mathematical inverse $\mathbf{G}^{-1}$.

3. There is a unique solution for **m**.

4. The data can be fit exactly.

6.11.3 Class II: *P* = *M* < *N*

Graphically, for this case, we have



1. $\mathbf{V}_0$ is empty since $P = M$.

2. $\mathbf{U}_0$ is not empty since $P < N$.

3. **G** has no mathematical inverse.

4. There is a unique solution in the sense that only one solution has the smallest misfit to the data.

5. The data cannot be fit exactly unless the *compatibility equations* are satisfied, which are defined as follows:

$$\underset{(N-P)\times N}{\mathbf{U}_0^{\mathrm{T}}} \quad \underset{N\times 1}{\mathbf{d}} \quad = \quad \underset{(N-P)\times 1}{\mathbf{0}} \tag{6.78}$$

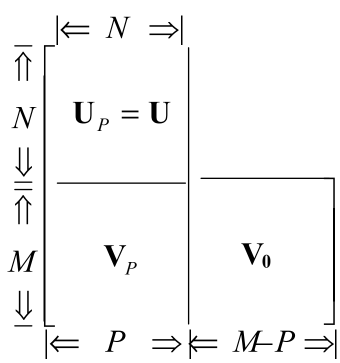
If the compatibility equations are satisfied, one can fit the data exactly. The compatibility equations are equivalent to saying that **d** has no projection onto $\mathbf{U}_0$.

Equation (6.78) can be thought of as the $N - P$ dot products of the eigenvectors in $\mathbf{U}_0$ with the data vector **d**. If all of the dot products are zero, then **d** has no component in the $(N - P)$-dimensional subset of N-space spanned by the vectors in $\mathbf{U}_0$. **G**, operating on any vector **m**, can only predict a vector that lies in the P-dimensional subset of N-space spanned by the P eigenvectors in $\mathbf{U}_P$. We will return to this later.

6. $P = M < N$ is the classic least squares environment. We will consider least squares again when we introduce the generalized inverse.

6.11.4 Class III: $P = N < M$

Graphically, for this case, we have



1. $\mathbf{U}_0$ is empty since $P = N$.

2. $\mathbf{V}_0$ is not empty since $P < M$.

3. **G** has no mathematical inverse.

4. You can fit the data exactly because $\mathbf{U}_0$ is empty.

5. Solution is not unique. If $\mathbf{m}^{\mathrm{est}}$ is a solution which fits the data exactly

$$\mathbf{Gm}^{\text{est}} = \mathbf{d}^{\text{pre}} = \mathbf{d} \tag{6.79}$$

then

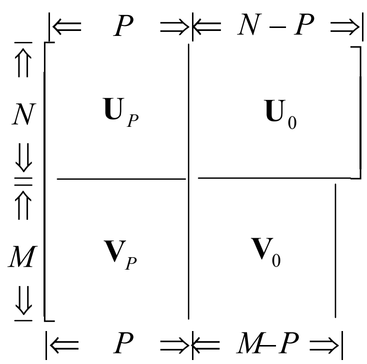
$$\mathbf{m}^{\text{est}} + \sum_{i=P+1}^{M} \alpha_i \mathbf{v}_i \tag{6.80}$$

is also a solution, where α_i is any arbitrary constant.

6. $P = N < M$ is the minimum length environment. The minimum length solution sets α_i, $i = (P + 1), \ldots, M$ to zero.

6.11.5 Class IV: $P < \min(N, M)$

Graphically, for this case, we have (next page)



1. Neither $\mathbf{U}_0$ nor $\mathbf{V}_0$ is empty.

2. **G** has no mathematical inverse.

3. You cannot fit the data exactly unless the compatibility equations (Equation 6.78) are satisfied.

4. The solution is nonunique.

This sounds like a pretty bleak environment. No mathematical inverse. Cannot fit the data. The solution is nonunique. It probably comes as no surprise that most realistic problems are of this type [$P < \min(N, M)$]!

In the next chapter we will introduce the *generalized inverse operator*. It will reduce to the unique mathematical inverse when $P = M = N$. It will reduce to the least squares operator when we have $P = M < N$, and to the minimum length operator when $P = N < M$. It will also give us a solution in the general case where we have $P < \min(N, M)$ that has many of the properties of the least squares and minimum length solutions.