Linear models with applications in R

PUBHLTH 744: Handout 4(Matrix Decompositions)

Instructor: Andrea S. Foulkes

Division of Biostatistics and Epidemiology
UMass School of Public Health and Health Sciences

Fall 2007

Matrix Decompositions

We consider 4 matrix decompositions relevant to linear modeling. Both the spectral decomposition and the Cholesky decomposition apply to $n \times n$ square matrices. The former requires the matrix be symmetric while the later assumes the matrix is positive definite. The QR factorization and the singular value decomposition (SVD) apply more broadly to $n \times p$ matrices.

- Singular Value Decomposition
- Spectral Decomposition
- QR Decomposition
- Cholesky Decomposition

Singular Value Decomposition (SVD)

Theorem 4.1: Singular Value Decomposition (SVD). For an $n \times p$ matrix A of rank r, there exists orthogonal matrices $U_{n\times p}$ and $V_{p\times p}$ such that

$$U'AV = \begin{pmatrix} \Delta & 0 \\ 0 & 0 \end{pmatrix}_{p\times p}$$

where $\Delta = diag(\delta_1, \ldots, \delta_r)$ and $\delta_1 \geq \ldots \geq \delta_r > 0$ are called the singular values of A.

Singular Value Decomposition (SVD)

Recall, for an orthogonal matrix X, we have $X^{-1} = X'$ or equivalently, $X'X = XX' = I$. Since U and V are orthogonal, we know $UU' = I$, $VV' = I$ and therefore $U'AV = D \Leftrightarrow UU'AVV' = UDV' \Leftrightarrow A = UDV'$.

Example

Consider the matrix $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Note that A is a 2×2 symmetric matrix of rank, $r(A) = 1$. The SVD is obtained using the `svd()` command as demonstrated below.

Singular Value Decomposition (SVD)

```
# Create the matrix A
> A <- matrix(c(1,1,1,1),2)
> A
     [,1] [,2]
[1,]    1    1
[2,]    1    1

# Calculate the singular value decomposition of A
> svd(A)
$d
[1] 2 0

$u
           [,1]       [,2]
[1,] -0.7071068 -0.7071068
[2,] -0.7071068  0.7071068

$v
           [,1]       [,2]
[1,] -0.7071068 -0.7071068
[2,] -0.7071068  0.7071068
```

Singular Value Decomposition (SVD)

- We know from the SVD theorem that we can write $D = U'AV$. Since A is of dimension 2×2 and $r(A) = 1$, we expect only one non-zero singular value, δ in the diagonal elements of D.
- We can also reconstruct A from the components of the SVD as shown below.

Singular Value Decomposition (SVD)

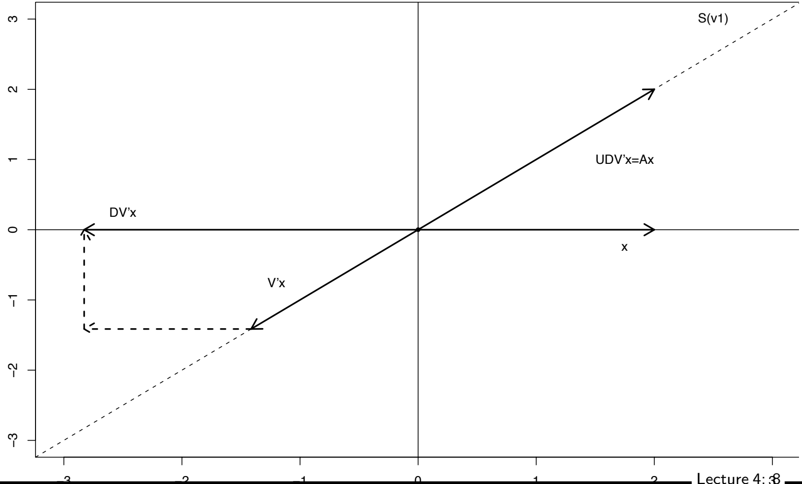
```
# Checking that U'AV=Delta
> V <- svd(A)$v
> U <- svd(A)$u
> round(t(U) %*% A %*% V,2)
    [,1] [,2]
[1,]    2    0
[2,]    0    0

# Generate A from the components of the SVD of A
> Delta <- diag(svd(A)$d)
> U %*% Delta %*% t(V)
     [,1] [,2]
[1,]    1    1
[2,]    1    1
```

Singular Value Decomposition (SVD)

Recall, matrices are spatial operators that rotate, stretch and shrink vectors. The SVD and the Spectral theorem tell us how to break up these operations of a matrix into three steps. Here we consider an example using the SVD, though a similar interpretation can be arrived at using the Spectral theorem. Let $A = UDV'$ act upon the vector $x = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$. It is easily seen that $Ax = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$. Now, let us consider $UDV'x$:

Singular Value Decomposition (SVD)



Singular Value Decomposition (SVD)

- First we have

$$x^* = V'x = \begin{pmatrix} -0.707 & -0.707 \\ -0.707 & 0.707 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} -1.414 \\ -1.414 \end{pmatrix}$$

- Note $V'x \in \mathcal{S}(v_1)$ where v_1 is an eigenvector of $A'A$. That is, pre-multiplying by the matrix V' rotates the vector x to be in the span of the first eigenvector of $A'A$.

Singular Value Decomposition (SVD)

- Since D is a diagonal matrix, pre-multiplying by D has the effect of stretching or shrinking the corresponding axes. In this example the first diagonal element equals 2 and the second equals 0 and so the effect is to stretch the x-axis by 2 and reduce the y axis to 0:

$$x^{**} = Dx^{*} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -1.414 \\ -1.414 \end{pmatrix} = \begin{pmatrix} -2.828 \\ 0 \end{pmatrix}$$

Singular Value Decomposition (SVD)

- Note that 2 and 0 are the square roots of the eigenvalues of $A'A$.
- Finally, pre-multiplying by U gives

$$Ux^{**} = \begin{pmatrix} -0.707 & -0.707 \\ -0.707 & 0.707 \end{pmatrix} \begin{pmatrix} -2.828 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

This has the effect of rotating x^{**} back to the original space.

Singular Value Decomposition (SVD)

- Also of note, the columns of U and V are the eigenvectors corresponding to the nonzero eigenvalues of AA' and $A'A$ respectively, and for $r(A) = r$, $\delta_1^2, \ldots, \delta_r^2$ are the eigenvalues of $A'A$.
- Finally, it can be shown that $A = UDV' = U_1 \Delta V_1'$ where U_1 is $n \times r$ and V_1 is $p \times r$, representing the orthonormal columns of U and V respectively that correspond to the singular values.

Spectral Decomposition

- If A is a symmetric, square matrix, then we know: (1) $A'A = AA'$ and (2) the eigenvalues of $A'A$ equal the squares of the eigenvalues of A. Proof of this second property is straightforward. Suppose v is an eigenvector of A with eigenvalue λ. Then

$$Av = \lambda v$$

or equivalently

$$A'Av = A'\lambda v = A\lambda v = \lambda(Av) = \lambda(\lambda v) = \lambda^2 v$$

or λ^2 is an eigenvalue of $A'A$. Note that we also have that the unit length eigenvector of A equals the unit length eigenvector of $A'A$.

Spectral Decomposition

- The first result, $A'A = AA'$ implies $U = V$. This follows from the facts that the columns of U and the columns of V are eigenvectors for the same matrices ($A'A = AA'$) and the columns of both U and V are of length 1. These results lead us to the following, related theorem regarding symmetric matrices.

Theorem 4.2: Spectral Decomposition. For an $n \times n$ symmetric matrix A, there exists an orthogonal matrix P such that $A = P\Lambda P'$ where $\Lambda = diag(\lambda_1, \ldots, \lambda_n)$, $\lambda_1 < \ldots < \lambda_n$ and P is the corresponding orthogonal matrix of eigenvectors.

Spectral Decomposition

Note that the Spectral decomposition is a special case of the SVD decomposition in which $U = V = P$ and $\lambda_i = \sqrt{\delta_i}$ for $i = 1, \ldots, n$. Both of these conditions hold in the case that A is symmetric.

Example

Again consider the matrix $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Since A is symmetric, we can use the spectral theorem to arrive at the the same results. That is, $U = V$ and $\delta_1 = \lambda_1^2$. Furthermore, U and V are orthogonal matrices with columns equal to the eigenvectors of A. That is, $U = V = P$ where P is defined in the Spectral theorem above. These are illustrated below.

Spectral Decomposition

```
# Check that U=V=P and \delta = \lambda^2
> round(U,2)==round(V,2)
     [,1] [,2]
[1,] TRUE TRUE
[2,] TRUE TRUE
> eigen(A)
$values
[1] 2.000000e+00 9.860761e-32
$vectors
          [,1]       [,2]
[1,] 0.7071068  0.7071068
[2,] 0.7071068 -0.7071068
> eigen(t(A)%*%A)
$values
[1] 4.000000e+00 1.972152e-31
$vectors
          [,1]       [,2]
[1,] 0.7071068  0.7071068
[2,] 0.7071068 -0.7071068
> eigen(A %*% t(A))
$values
[1] 4.000000e+00 1.972152e-31
$vectors
          [,1]       [,2]
[1,] 0.7071068  0.7071068
[2,] 0.7071068 -0.7071068
```

QR Factorization

Theorem 4.3: QR Factorization. Suppose A is an $n \times p$ matrix, then A can be written in the form $A = QR$ where $Q_{n\times p}$ has orthnormal columns and $R_{p\times p}$ is an upper triangular matrix. If the columns of A are linearly independent (i.e. A is non-singular), then this decomposition is unique and R has positive diagonal elements.

QR Factorization

Example Let $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. Since the columns of A are linearly independent we know by the QR factorization theorem that we can write $A = QR$ where Q has orthonormal columns and R is an upper triangular matrix with positive diagonal elements. Calculation of these matrices is achieved using the `qr()`, `qr.Q()` and `qr.R()` functions in R as shown below. Note that both the `qr.Q()` and `qr.R()` functions require an object resulting from a QR decomposition, as from a call to `qr`.

QR Factorization

```
# Define A
> A <- matrix(c(1,1,1,2),2)
> A
     [,1] [,2]
[1,]    1    1
[2,]    1    2

# Calculate the Q matrix from the QR factorization of A
> Q <- qr.Q(qr(A))
> Q
           [,1]       [,2]
[1,] -0.7071068 -0.7071068
[2,] -0.7071068  0.7071068

# Calculate the R matrix from the QR factorization of A
> R <- qr.R(qr(A))
> R
          [,1]       [,2]
[1,] -1.414214 -2.1213203
[2,]  0.000000  0.7071068
```

QR Factorization

As expected, Q has orthonormal columns and R is an upper triangular matrix with positive diagonal elements. We confirm this, and that we can reconstruct A using Q and R with the following code:

```
# Checking Q is orthogonal
> Q%*%Q
               [,1]          [,2]
[1,]  1.000000e+00 -7.852334e-17
[2,] -7.852334e-17  1.000000e+00

# Reconstructing A from Q and R
> Q%*%R
     [,1] [,2]
[1,]    1    1
[2,]    1    2
```

QR Factorization

Notably, the `lsfit()` function in R also returns an object that can be used with the `qr.Q()` and `qr.R()` commands. We return to this later. The QR decomposition also provides an efficient way of calculating the determinant of a matrix and the product of its eigenvalues. It can be shown that $det(A) = det(Q) * det(R)$. Furthermore $|det(Q)| = 1$ and so $|det(A)| = |det(R)|$. But the determinant of R is given by $det(R) = \prod_i r_{ii}$, the product of the diagonal elements of R since R is triangular. We saw previously, that the determinant of a matrix is equal to the product of its eigenvalues and so $|\prod_i r_{ii}| = |\prod_i \lambda_i|$

QR Factorization

Finally, the `qr()` function in R also returns the rank of a matrix. This is demonstrated below.

```
# Using the qr() function to calculate the rank of A
> qr(A)$rank
[1] 2
```

Cholesky Decomposition

Theorem 4.4: Cholesky Decomposition. A square and positive definite (definition given below) matrix, A can be written $A = LL'$ where L is a lower triangular matrix with strictly positive diagonal elements.

Cholesky Decomposition

Example Define A as in example above. We know A is symmetric. A is also positive definite since the eigenvalues of A are all strictly greater than 0:

```
# Check the eigenvalues of $A$ are greater than 0
> eigen(A)
$values
[1] 2.618034 0.381966
$vectors
          [,1]       [,2]
[1,] 0.5257311  0.8506508
[2,] 0.8506508 -0.5257311
# We could also use the following code to check the
# eigenvalues are greater than 0
> sum(eigen(A)$values<1e-8)
[1] 0
# Confirm A is symmetric
> sum(t(A)!=A)
[1] 0
```

Cholesky Decomposition

As a result, we can apply the Cholesky decomposition.

```
# Determining the Cholesky decomposition of A
 > chol(A)
     [,1] [,2]
[1,]    1    1
[2,]    0    1
```

Cholesky Decomposition

Only a single matrix is returned since the Cholesky decomposition tells us A is the product of a lower triangular matrix and its transpose, $A = LL'$. The matrix returned by R is the upper triangular matrix, L'. As a result, we can reconstruct A as follows:

```
# Reconstructing A from the output of chol()
> Lt <- chol(A)
> t(Lt)%*%Lt
     [,1] [,2]
[1,]    1    1
[2,]    1    2
```