

# Modelos matemáticos y numéricos

Proyecto #1: Ejercicios DVS

Alumno: José Manuel Proudinat Silva

## 1. Algoritmo para DVS

En esta primera sección tenemos como objetivo definir el algoritmo más difundido para la Descomposición en Valores Singulares (DVS) que sea numéricamente estable, además de mostrar la complejidad del mismo.

El algoritmo más utilizado para realizar esta descomposición de matrices (el cual está implementado en Matlab y Lapack) es una versión revisada del algoritmo presentado por Golub y Van Loan. Esta versión revisada fue presentada por J. Demmel y W. Kahan en su artículo "Accurate Singular Values of Bidiagonal Matrices" (1990). La idea detrás del mismo la podemos ver en dos pasos: Primero, la matriz se reduce a una forma bidiagonal  $B$  a través de transformaciones de Householder. Durante estas iteraciones podemos construir  $U$  y  $V$  multiplicando las matrices de rotación que surgen mientras iteramos. De esta forma llegamos a una descomposición de la forma  $A = UBV^T$ .

Finalmente comenzamos un proceso iterativo con la matriz bidiagonal en el cual vamos haciendo ceros los valores de  $B$  que cumplan que  $|b_{i,i+1}| \leq \epsilon(|b_{i,i}| + |b_{i+1,i+1}|)$ . Partimos  $B$  en 3 bloques, tal que el tercer bloque sea diagonal y el segundo no tenga ceros en la diagonal superior. Si encontramos un cero en la diagonal del segundo bloque se aplican rotaciones de Givens, de tal forma que el segundo bloque siga siendo bidiagonal superior (Golub-Reinsch). Si no hay ceros, aplicamos otro subalgoritmo con rotaciones de Givens que van haciendo más pequeños los valores fuera de la diagonal de  $B$  (Golub-Kahan). Iteramos hasta que  $B$  sea una matriz completamente diagonal.

Golub y Van Loan muestran en su libro "Matrix computation" que este algoritmo es de orden  $\mathcal{O}(m^2n + n^3)$  si queremos hallar la descomposición completa.

## 2. Problema de mínimos cuadrados

Ahora mostraremos con argumentos de álgebra lineal como el problema de mínimos cuadrados  $Ax = b$  tiene solución en  $\bar{x} = (A^T A)^{-1} A^T b$ . Donde  $A$  es una matriz de  $m \times n$  con  $m \geq n$  de rango completo.

Como  $A$  es de rango completo, entonces  $A^T A$  es simétrica positiva definida, y por lo tanto es invertible.

Buscamos  $\bar{x}$  tal que minimice el error dado por  $r = b - A\bar{x}$ . El cual en Álgebra Lineal se demuestra que está dada por la proyección en  $A$  de  $b$ . Por lo cual tomando  $A\bar{x}$  como la proyección, tenemos que esta debe ser ortogonal al residuo. Es decir,  $(b - A\bar{x})^T A\bar{x} = 0$ . Con lo que llegamos a lo siguiente:

$$\begin{aligned}
& (A\bar{x})^T(b - A\bar{x}) = 0 \\
& \Leftrightarrow \bar{x}^T A^T b - \bar{x}^T A^T A \bar{x} = 0 \\
& \Leftrightarrow \bar{x}^T (A^T b - A^T A \bar{x}) = 0 \\
& \Leftrightarrow A^T b - A^T A \bar{x} = 0 \\
& \Leftrightarrow A^T A \bar{x} = A^T b \\
& \Leftrightarrow \bar{x} = (A^T A)^{-1} A^T b
\end{aligned}$$

### 3. Componentes principales y DVS

A continuación haremos una interpretación del Análisis de Componentes Principales (ACP) tomando la idea de la Descomposición en Valores Singulares.

Recordemos que el ACP requiere que se calculen los valores propios y vectores propios de su matriz de covarianzas  $XX^T$ , donde  $X$  es la matriz de datos. Como la matriz de covarianzas es simétrica podemos hacer la descomposición espectral de la matriz:

$$XX^T = QDQ^T$$

Donde  $D$  es la matriz diagonal de valores propios y  $Q$  es una matriz ortogonal conteniedo a los vectores propios.

Usando la DVS de  $X$  podemos llegar a que:

$$\begin{aligned}
XX^T &= (U\Sigma V^T)(U\Sigma V^T)^T \\
&= U\Sigma V^T V \Sigma U^T \\
&= U\Sigma^2 U^T
\end{aligned}$$

De esta manera la interpretación es sencilla. Observamos que los cuadrados de los valores singulares de  $X$  contenidos en  $\Sigma$  representan la varianza que logra representar la combinación lineal de los datos dada por los vectores en  $U$ . Así observamos también que los vectores de  $U$  representan las combinaciones lineales que nos dan las distintas reducciones de dimensiones que buscan representar lo máximo posible de la varianza total de los datos.

### 4. Factorización de Schur

La descomposición (factorización) de Schur de una matriz cuadrada  $A$ , cuyos elementos pueden ser complejos, es una descomposición de la forma:

$$Q^H A Q = T = D + N$$

donde  $Q$  es una matriz unitaria, es decir  $Q^{-1} = Q^H$ , donde  $Q^H$  es la matriz conjugada transpuesta de  $Q$ .  $T$  es una matriz triangular superior la cual puede ser expresada como la suma de

$D = \text{diag}(\lambda_1, \dots, \lambda_n)$  (donde los  $\lambda_i$  son los valores propios de la matriz  $A$ ) más una matriz triangular superior estricta  $N$ .

El impacto práctico de esta descomposición es que lo único que requiere de la matriz  $A$  es que sea cuadrada para poder aplicar la descomposición, además de que podemos obtener los valores propios de la matriz  $A$  en la matriz diagonal  $D$ . También notamos que si la matriz  $A$  es real, entonces la factorización es de la forma  $Q^T A Q = T = D + N$ . Con lo que fácilmente podemos llegar a la descomposición espectral de la matriz haciendo uso de Householder para eliminar los valores por encima de la diagonal.

## Referencias

- [1] Gene H. Golub Charles F. Van Loan, *Matrix computations*. 3th edition. Josh Hopkins University. 1996.
- [2] Alan Kaylor Cline Inderjit S. Dhillon, *Computation of the Singular Value Decomposition. Handbook of Linear Algebra*. Discrete mathematics and its application. Series editor KENNETH H. ROSEN. Chapman Hall/CRC.2007.