

Guía N^o 1 Computación Científica I Valparaíso–Santiago, 3 de abril de 2008

ÍNDICE

1. Normas de vectores y matrices.	4
2. Descomposición de Valor Singular	11
3. Mis Dudas Castellanas	20
Referencias	22

I. Ficha Técnica de la Asignatura.

1. PROGRAMA.

ASIGNATURA:	Computación Científica I.
SIGLA:	ILI-285.
CRÉDITOS:	4.
PRERREQUISITO:	MAT-023 Matemática III.
EXAMEN:	No tiene.
HORAS DE CÁTEDRA SEMANALES:	4.
HORAS DE AYUDANTÍA SEMANALES:	2 (en promedio).
HORAS DE LABORATORIO SEMANALES:	2 (en promedio).

Objetivos: Al aprobar la asignatura el alumno *debería ser capaz* de:

- (1) Seleccionar las estructuras de datos, los algoritmos y las plataformas computacionales más apropiadas para abordar problemas concretos que pueden reducirse al modelado computacional lineal.
- (2) Resolver los problemas mencionados, teniendo en cuenta la complejidad y el costo de los algoritmos involucrados.
- (3) Integrar técnicas de programación y modelado computacional lineal en aplicaciones de Matemáticas, Ciencias Naturales y Ciencias de la Ingeniería.
- (4) Emplear programas de modelado computacional lineal de uso más frecuente en la Ingeniería (por ejemplo, Matlab(R), Mathematica(R), Maple(R)).

Contenido:

- (1) Introducción a la computación científica y a la computación de alto rendimiento. Elementos a considerar en cada ítem: plataformas computacionales para los métodos cuantitativos; estructuras de datos más afines; complejidad de los algoritmos involucrados.
- (2) Sistemas lineales en general. Computación matricial. Computación con matrices ralas. Condicionamiento y estabilidad. Estructuras de datos más afines y complejidad de los algoritmos involucrados.
- (3) Problemas computacionales de valores y vectores propios. Formas normales de Jordan, Frobenius y Schur. Elementos de la teoría espectral. Estructuras de datos más afines y complejidad de los algoritmos involucrados.
- (4) Métodos iterativos y sistemas lineales de gran escala. Elementos de programación lineal. Modelos computacionales lineales y elementos de la programación paralela.
- (5) Elementos de la teoría y control de errores en el cálculo computacional lineal.
- (6) Aplicaciones a la Estadística y la Ingeniería. Mínimos cuadrados. Multicolinealidad. Matrices de correlación. Teoría de redes. Elementos de la teoría de juegos.

Bibliografía ampliada:

- [1] A.V. Aho and J.E. Hopcroft and J.D. Ullman. *The design and analysis of computer algorithms*. Addison-Wesley, Reading, Massachusetts, 1974.
- [2] D.J. Higham, N.J. Higham. *Matlab Guide*. Society for Industrial & Applied Mathematics, Philadelphia, PA, 2000.
- [3] Seymour Lipschutz. *3,000 Solved Problems in Linear Algebra*. McGraw-Hill, Inc., New York, NY 10121-2298, 1989. This is a good source for elementary exercises!
- [4] Seymour Lipschutz, Marc Lipson. *Schaum's Outline of Linear Algebra*. McGraw-Hill, Inc., New York, NY 10121-2298, 2000. This is a good source for elementary theory and exercises!
- [5] Seymour Lipschutz, Marc Lipson. *Schaum's Easy Outline of Linear Algebra* McGraw-Hill, Inc., New York, NY 10121-2298, 2002. Also a good source for elementary theory and exercises!
- [6] G. Lindfield and J. Penny. *Numerical Methods using MATLAB*. Prentice Hall, Upper Saddle River, NJ 07458, 2000.
- [7] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, Cambridge, 1993.
- [8] A. Quarteroni and R. Saco and F. Saleri. *Numerische Mathematik 1 und 2* Springer Verlag, Berlin-Heidelberg, 2002.
- [9] J. Stoer, U. Bulirsch. *Introduction to Numerical Analysis*. Springer Verlag, New York - Berlin, 1992.
- [10] G. Strang. *Introduction to Linear Algebra*. Wellesley-Cambridge Press, Wellesley MA 02181 USA, 1993. *Note*: A very good elementary book indeed!
- [11] L.N. Trefethen and D. Bau III. *Numerical Linear Algebra*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1997. *Note*: This is **the** book. A very good book indeed! More advanced than Strang's book.
- [12] L.N. Trefethen. *Spectral Methods in MATLAB*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2000. This is "the" text!

2. DE LOS PRERREQUISITOS: ILI-285 es la primera de una *secuencia* de dos asignaturas: ILI-285 e ILI-286. El prerrequisito canónico de esta *secuencia* de asignaturas es MAT-024 Matemática IV.

Sin embargo, con el objeto de no retrasar el avance académico de los alumnos que deben o desean cursar ILI-285 *sin* haber aprobado todavía el prerrequisito MAT-024, se ha decidido aceptar *excepcionalmente* que inscriban la asignatura ILI-285 en el primer semestre 2008, siempre y cuando tengan aprobada la asignatura MAT-023.

Para inscribir ILI-286 en el segundo semestre 2008 se *exigirá* haber aprobado tanto MAT-024 como ILI-285.

3. DE LA EVALUACIÓN: La asignatura considera los siguientes instrumentos de evaluación:

- (1) Certamen de Cátedra N^o 1: Sábado 12.04.2008, 10:00 horas.
Origina una nota *individual* N_1 .
- (2) Certamen de Cátedra N^o 2: Sábado 17.05.2008, 10:00 horas.
Origina una nota *individual* N_2 .
- (3) Certamen de Cátedra N^o 3: Sábado 21.06.2008, 10:00 horas.
Origina una nota *individual* N_3 .
- (4) Las fechas señaladas son provisorias y deberán ser precisadas durante el transcurso del semestre.
- (5) El alumno que **no** rinda uno o más de los certámenes regulares de cátedra, por las razones que fuere, recibirá la calificación *provisoria* $N_j = 0$ en esos certámenes.
- (6) Habrá además un **Certamen Recuperativo Voluntario** abierto a todos los alumnos. Fecha de este certamen: Sábado 28.06.2008, 10:00 horas. La fecha señalada es provisorias y deberá ser precisada durante el transcurso del semestre.
 - El certamen recuperativo tendrá tres secciones, correspondientes cada una a las materias cubiertas por los tres certámenes regulares de cátedra.

- El alumno que, habiendo rendido todos los certámenes regulares de cátedra y rindiere el Certamen Recuperativo, deberá abordar aquella sección en la que hubiere obtenido la menor calificación.

El alumno que, por las razones que fuere, no hubiere rendido uno o dos certámenes de cátedra, pero no más, deberá abordar la o las secciones correspondientes a los certámenes no rendidos.

- El estudiante recibirá una nota N'_j en cada sección j que aborde.
- Las notas N'_j sustituirán a las notas N_j de los certámenes regulares de cátedra.
- No habrá certamen recuperativo del certamen recuperativo (!).

Obsérvese que esta regulación interna de la asignatura cubre el caso de los alumnos que, por la razón que fuere, no hayan rendido algunos de los certámenes regulares de cátedra.

- (7) A lo largo del semestre se propondrá un abundante número de ejercicios teóricos y prácticos (trabajos computacionales de laboratorio) que los alumnos deberán abordar (i.e., estudiar, resolver, implementar, discutir, etc.) en las Ayudantías y los Laboratorios de la Asignatura. *El trabajo de los estudiantes en la solución de estos ejercicios y trabajos de laboratorio será evaluado mediante tres (3) controles individuales.*

El promedio aritmético de las calificaciones obtenidas en esos tres *controles*, dará origen a una nota *individual* N_4 para cada estudiante.

- (8) Desarrollo de un tema de investigación (mayormente bibliográfica) relacionado con el temario de la asignatura, especialmente acerca de las aplicaciones de la Computación Científica en la Informática, la Ciencia de la Computación y las diversas especialidades de la Ingeniería y la Ciencia.

Cada tema debe ser abordado por un equipo formado por unos diez (10) alumnos. Los temas específicos podrán ser propuestos por los propios integrantes del grupo, o bien asignados por los Profesores y Ayudantes. La proposición de temas y la composición de los grupos debe ser informada por escrito a los Ayudantes antes del 30 de Abril del año en curso.

Los avances y principales resultados de esta actividad deberán ser expuestos periódicamente a lo largo del semestre, en una o más conferencias de no más de 15 minutos, dictadas por uno de los miembros del grupo correspondiente, elegido al azar. Las conferencias incluyen un período de preguntas que deberán ser respondidas por cualquiera de los miembros del equipo de trabajo. Cada una de estas conferencias recibirá una calificación. El promedio de estas calificaciones dará origen a una nota N'_5 .

Adicionalmente, cada grupo deberá presentar un informe final relativo al tema estudiado y deberá presentar en ayudantía la implementación computacional de un algoritmo relevante en el tema estudiado. Esta actividad dará origen a una nota N''_5 .

El promedio de las notas N'_5 y N''_5 constituirá una nota N_5 que será la misma para todos los integrantes del equipo de trabajo.

- (9) La nota final N_F se calculará mediante la siguiente fórmula:

$$N_F = \begin{cases} \frac{1}{5} \sum_{j=1}^5 N_j, & \text{si } \frac{1}{3} \sum_{j=1}^3 N_j \geq 50, \\ \frac{1}{3} \sum_{j=1}^3 N_j, & \text{si } \frac{1}{3} \sum_{j=1}^3 N_j < 50. \end{cases}$$

- (10) Cualquier situación no considerada en este reglamento interno será resuelta por los Profesores de la asignatura.

Nota: Un problema abierto en la evaluación cuantitativa de cualquier actividad, incluyendo los procesos de enseñanza-aprendizaje, es la determinación de ciertos *parámetros descriptores*, justos y equitativos, que representen razonablemente bien los logros alcanzados en esa actividad.

En los procesos lectivos, por lo general, los diversos parámetros descriptores finalmente se reducen a uno solo, N_F , mediante alguna fórmula más o menos arbitraria, v.g., promedio geométrico, promedio

aritmético, promedio ponderado, etc. A veces, para efectos de decidir la aprobación o reprobación de una asignatura, en lugar de considerar un solo valor NF , es preferible considerar un conjunto de parámetros descriptores significativos y *regiones admisibles* para sus valores.

Un excelente ejercicio de modelación matemática y de Computación Científica, lo constituye la determinación de un esquema de evaluación que permita estimar adecuadamente los logros académicos de los estudiantes en una asignatura.

Los alumnos quedan cordialmente invitados a reflexionar sobre este tema y a plantear fundamentalmente sus soluciones mediante artículos *ad hoc*.

El equipo docente de ILI-285 le desea una provechosa estada en la asignatura, mucho éxito en su trabajo, y le recuerda que estará siempre atento a recibir sus comentarios, consultas, sugerencias y observaciones.

1. NORMAS DE VECTORES Y MATRICES.

1. *Las p-normas de vectores.* Consideremos el espacio vectorial \mathbb{C}^m . Para $x = [x_1, \dots, x_m]^T \in \mathbb{C}^m$ y $1 \leq p \in \mathbb{R} \cup \{\infty\}$ se define:

$$\|x\|_p := \begin{cases} \left(\sum_{k=1}^m |x_k|^p \right)^{1/p}, & \text{si } 1 \leq p < \infty, \\ \max_{1 \leq k \leq m} |x_k|, & \text{si } p = \infty. \end{cases} \quad (1)$$

Evidentemente, $\|x\|_p \geq 0$ y $\|\lambda x\|_p = |\lambda| \|x\|_p$ para todo $\lambda \in \mathbb{C}$, todo $x \in \mathbb{C}^m$ y todo $1 \leq p \in \mathbb{R} \cup \{\infty\}$. Sean $x, y \in \mathbb{C}^m$. Para $p = 1$, por la desigualdad triangular en \mathbb{C} , se tiene:

$$\|x + y\|_1 = \sum_{k=1}^m (|x_k + y_k|) \leq \sum_{k=1}^m |x_k| + \sum_{k=1}^m |y_k| = \|x\|_1 + \|y\|_1. \quad (2)$$

Por su parte, para $p = \infty$ se tiene:

$$\|x + y\|_\infty = \max_{1 \leq k \leq m} |x_k + y_k| \leq \max_{1 \leq k \leq m} |x_k| + \max_{1 \leq k \leq m} |y_k| = \|x\|_\infty + \|y\|_\infty. \quad (3)$$

Luego, $\|\cdot\|_1$ y $\|\cdot\|_\infty$ satisfacen la desigualdad triangular y, por lo tanto, son normas en \mathbb{C}^m . En lo que sigue se demostrará que $\|\cdot\|_p$ satisface la desigualdad triangular para todo $p \in \mathbb{R}$ con $p > 1$.

Se dice que $p, q \in [1, \infty]$ son *mutuamente conjugados* ssi

$$\frac{1}{p} + \frac{1}{q} = 1. \quad (4)$$

Notar que $p = 1$ y $q = \infty$ son mutuamente conjugados.

El siguiente es un resultado auxiliar que permite obtener con relativa facilidad la desigualdad triangular para $1 < p < \infty$.

(a) Sean $0 \leq u, v \in \mathbb{R}$ y $1 < p, q \in \mathbb{R}$ mutuamente conjugados, i.e., $(1/p) + (1/q) = 1$. Entonces:

$$uv = \inf \left\{ \frac{1}{p} t^p u^p + \frac{1}{q} t^{-q} v^q : t > 0 \right\}. \quad (5)$$

Demostración. ¡Tarea! ■

(b) *La desigualdad de Hölder.* Sean $p, q \in [1, \infty]$ mutuamente conjugados. Entonces se tiene:

$$|x^* y| = \left| \sum_{k=1}^m x_k y_k \right| \leq \sum_{k=1}^m |x_k y_k| \leq \|x\|_p \|y\|_q, \quad \text{para todo } x, y \in \mathbb{C}^m. \quad (6)$$

Nótese que para $p = 2$ la desigualdad de Hölder se reduce a la conocida desigualdad de Cauchy-Schwartz.

Demostración. Sean $x, y \in \mathbb{C}^m$. Para $p = 1$ y $q = \infty$ se tiene:

$$\sum_{k=1}^m |x_k y_k| \leq \sum_{k=1}^m |x_k| \max_{1 \leq s \leq m} |y_s| = \left(\sum_{k=1}^m |x_k| \right) \max_{1 \leq s \leq m} |y_s| = \|x\|_1 \|y\|_\infty,$$

lo que prueba la desigualdad de Hölder para estos valores de p y q .

Ahora consideramos el caso $1 < p, q < \infty$ mutuamente conjugados. Sobre la base del resultado (a), con $u = |x_k|$ y $v = |y_k|$, se tiene:

$$|x_k y_k| \leq \frac{1}{p} t^p |x_k|^p + \frac{1}{q} t^{-q} |y_k|^q, \quad t > 0, \quad k = 1 : m.$$

Sumando estas m desigualdades resulta:

$$\sum_{k=1}^m |x_k y_k| \leq \frac{1}{p} t^p \sum_{k=1}^m |x_k|^p + \frac{1}{q} t^{-q} \sum_{k=1}^m |y_k|^q, \quad t > 0.$$

Tomando el *infimum* sobre $t > 0$ en el lado derecho de la desigualdad precedente y en vista del resultado (a) — con $u^p = \sum_{k=1}^m |x_k|^p$ y $v^q = \sum_{k=1}^m |y_k|^q$ — se obtiene:

$$\sum_{k=1}^m |x_k y_k| \leq \inf \left\{ \frac{1}{p} t^p \sum_{k=1}^m |x_k|^p + \frac{1}{q} t^{-q} \sum_{k=1}^m |y_k|^q : t > 0 \right\} = \left(\sum_{k=1}^m |x_k|^p \right)^{1/p} \left(\sum_{k=1}^m |y_k|^q \right)^{1/q},$$

lo que concluye la demostración. ■

(c) *La desigualdad de Minkowski.* Para todo $p, q \in [1, \infty]$ se tiene:

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p, \quad \text{para todo } x, y \in \mathbb{C}^m. \quad (7)$$

Demostración. Los casos $p = 1$ y $p = \infty$ ya fueron verificados en (2) y (3) respectivamente. Para $x + y = 0$ la desigualdad es evidente. Luego, sin pérdida de generalidad (s.P.d.G.) se puede suponer $1 < p < \infty$ y $x + y \neq 0$. Defínase $q := p/(p-1)$, de modo que se cumple $(1/p) + (1/q) = 1$. Aplicando la desigualdad de Hölder se deduce entonces:

$$\begin{aligned} \sum_{k=1}^m |x_k + y_k|^p &= \sum_{k=1}^m |x_k + y_k| |x_k + y_k|^{p-1} \\ &\leq \sum_{k=1}^m |x_k| |x_k + y_k|^{p-1} + \sum_{k=1}^m |y_k| |x_k + y_k|^{p-1} \\ &\leq \left(\sum_{k=1}^m |x_k|^p \right)^{1/p} \left(\sum_{k=1}^m |x_k + y_k|^{(p-1)q} \right)^{1/q} \\ &\quad + \left(\sum_{k=1}^m |y_k|^p \right)^{1/p} \left(\sum_{k=1}^m |x_k + y_k|^{(p-1)q} \right)^{1/q} \\ &= \left[\left(\sum_{k=1}^m |x_k|^p \right)^{1/p} + \left(\sum_{k=1}^m |y_k|^p \right)^{1/p} \right] \left(\sum_{k=1}^m |x_k + y_k|^p \right)^{(p-1)/p} \\ &= \left[\left(\sum_{k=1}^m |x_k|^p \right)^{1/p} + \left(\sum_{k=1}^m |y_k|^p \right)^{1/p} \right] \left(\sum_{k=1}^m |x_k + y_k|^p \right)^{1/q}, \end{aligned}$$

donde se ha utilizado la identidad:

$$\frac{p-1}{p} = 1 - \frac{1}{p} = \frac{1}{q}.$$

Considerando el primero y el último término de la cadena de desigualdades precedentes y, puesto que $x + y \neq 0$, dividiendo por $(\sum_{k=1}^m |x_k + y_k|^p)^{(p-1)/p}$ se obtiene:

$$\left(\sum_{k=1}^m |x_k + y_k|^p\right)^{1/p} = \left(\sum_{k=1}^m |x_k + y_k|^p\right)^{1-1/q} \leq \left(\sum_{k=1}^m |x_k|^p\right)^{1/p} + \left(\sum_{k=1}^m |y_k|^p\right)^{1/p},$$

que es la desigualdad triangular planteada. ■

2. [2, §5.6, p. 290] *Normas matriciales.* Se dice que una función $\|\cdot\| : M(m \times n, \mathbb{C}) \rightarrow \mathbb{R}$ es una *norma matricial* si y sólo si para todo $A, B \in M(m \times n, \mathbb{C})$ se cumplen los siguientes axiomas:

- (i) $\|A\| \geq 0$ y $\|A\| = 0$ si y sólo si $A = 0$,
- (ii) $\|\lambda A\| = |\lambda| \|A\|$ para todo $\lambda \in \mathbb{C}$,
- (iii) $\|A + B\| \leq \|A\| + \|B\|$,
- (iv) $\|AB\| \leq \|A\| \|B\|$.

Los axiomas (i)-(iii) expresan simplemente que las normas matriciales deben ser, en particular, normas vectoriales sobre $M(m \times n, \mathbb{C})$ considerado como espacio vectorial. Algunos autores, como G.H. Golub y C.F. Van Loan [1, (2.3.4), p. 55], no exigen el axioma (iv). Golub y Van Loan aclaran, sin embargo, que por lo general se trabaja con normas matriciales que sí satisfacen el axioma (iv).

En vista del axioma (iv), para toda matriz no nula A con $A^2 = A$ se tiene $\|A^2\| = \|AA\| \leq \|A\| \|A\| = \|A\|^2$ y, por lo tanto, $\|A\| \geq 1$. En particular, $\|I\| \geq 1$.

Si A es invertible, se tiene $I = AA^{-1}$ y el axioma (iv) implica $\|I\| = \|AA^{-1}\| \leq \|A\| \|A^{-1}\|$ y en consecuencia $\|A^{-1}\| \geq \|I\| / \|A\|$.

Sea $A = (a_{ij}) \in M(m \times n, \mathbb{C})$. Las siguientes son algunas normas matriciales populares:

2.1 La l_1 de la matriz A considerada como un vector en \mathbb{C}^{mn} :

$$\|A\|_{l_1} = \sum_{i=1}^m \sum_{j=1}^n |a_{ij}| = \sum_{j=1}^n \|A^j\|_1,$$

donde A^j denota la j -ésima columna de la matriz A .

2.2 La norma de *Frobenius* o norma l_2 de la matriz A considerada como un vector en \mathbb{C}^{mn} :

$$\|A\|_F \equiv \|A\|_{l_2} = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2\right)^{1/2} = \left(\sum_{j=1}^n \|A^j\|_2^2\right)^{1/2} = (\text{tr}(A^*A))^{1/2}$$

También se le conoce como norma de *Schur* o norma de *Hilbert-Schmidt*.

2.3 La norma-1 de la matriz A se define como el *máximo de las normas l_1 de las columnas*:

$$\|A\|_1 = \max_{1 \leq j \leq n} \|A^j\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|,$$

donde A^j denota la j -ésima columna de A .

2.4 La norma- ∞ de la matriz A se define como el *máximo de las normas l_1 de las filas*:

$$\|A\|_\infty = \max_{1 \leq i \leq m} \|A_i\|_1 = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|,$$

donde A_i denota la i -ésima fila de A .

TAREA: Verifique que las precedentes son efectivamente normas de matrices que poseen las propiedades que se indican.

Desarrollo de (3). Sólo verificaremos *submultiplicatividad* de esta norma. Sean $A \in M(m \times p, \mathbb{C})$ y $B \in M(p \times n, \mathbb{C})$, de modo que $AB \in M(m \times n, \mathbb{C})$. Note que las matrices no son cuadradas pero el producto AB queda bien definido. Entonces se tiene:

$$\begin{aligned} \|AB\|_1 &= \max \left\| \left[|A \cdot B^1| |A \cdot B^2| \dots |A \cdot B^n| \right] \right\|_1 \\ &= \max \left\{ \|A \cdot B^1\|_1, \|A \cdot B^2\|_1, \dots, \|A \cdot B^n\|_1 \right\}. \end{aligned} \tag{8}$$

Ahora bien, para cada término $\|A \cdot B^j\|_1$ en (8) se tiene:

$$\begin{aligned} \|A \cdot B^j\|_1 &= \|b_{1,j}A^1 + b_{2,j}A^2 + \dots + b_{n,j}A^n\|_1 \\ &\leq |b_{1,j}| \|A^1\|_1 + |b_{2,j}| \|A^2\|_1 + \dots + |b_{n,j}| \|A^n\|_1 \\ &\leq \text{máx} \{ \|A^1\|_1, \|A^2\|_1, \dots, \|A^n\|_1 \} (|b_{1,j}| + |b_{2,j}| + \dots + |b_{n,j}|) \\ &= \|A\|_1 \|B^j\|_1. \end{aligned} \tag{9}$$

De (8) y (9) se obtiene finalmente:

$$\begin{aligned} \|AB\|_1 &\leq \text{máx} \{ \|A\|_1 \|B^1\|_1, \|A\|_1 \|B^2\|_1, \dots, \|A\|_1 \|B^n\|_1 \} \\ &= \|A\|_1 \text{máx} \{ \|B^1\|_1, \|B^2\|_1, \dots, \|B^n\|_1 \} \\ &= \|A\|_1 \|B\|_1. \end{aligned} \tag{10}$$

3. Como Ud. recordará la norma de Frobenius o de Hilbert-Schmidt de una matriz $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ con $a, b, c, d \in \mathbb{C}$, se define como $\|A\|_F = \sqrt{|a|^2 + |b|^2 + |c|^2 + |d|^2}$. Verifique que esta norma matricial posee la propiedad de la submultiplicatividad, i.e., $\|AB\|_F \leq \|A\|_F \|B\|_F, \forall A, B \in M(2 \times 2, \mathbb{C})$.

Solución: Sean $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$, con $a_{ij}, b_{ij} \in \mathbb{C}$. Escribamos $A_i = [a_{i,1}, a_{i,2}]^T \in \mathbb{C}^2, i = 1, 2$, y $B^j = [b_{1,j}, b_{2,j}]^T \in \mathbb{C}^2, j = 1, 2$. Entonces $AB = \begin{bmatrix} A_1^T B^1 & A_1^T B^2 \\ A_2^T B^1 & A_2^T B^2 \end{bmatrix}$. Aplicando la desigualdad de Cauchy-Schwarz usual para la norma $\|\cdot\|_2$ de \mathbb{C}^2 se obtiene:

$$\begin{aligned} \|AB\|_F^2 &= |A_1^T B^1|^2 + |A_1^T B^2|^2 + |A_2^T B^1|^2 + |A_2^T B^2|^2 \\ &\leq \|A_1\|_2^2 \|B^1\|_2^2 + \|A_1\|_2^2 \|B^2\|_2^2 + \|A_2\|_2^2 \|B^1\|_2^2 + \|A_2\|_2^2 \|B^2\|_2^2 \\ &= (\|A_1\|_2^2 + \|A_2\|_2^2) \cdot (\|B^1\|_2^2 + \|B^2\|_2^2) \\ &= (|a_{11}|^2 + |a_{12}|^2 + |a_{21}|^2 + |a_{22}|^2) \cdot (|b_{11}|^2 + |b_{21}|^2 + |b_{12}|^2 + |b_{22}|^2) \\ &= \|A\|_F^2 \|B\|_F^2, \end{aligned}$$

que era lo que se quería demostrar.

4. Sea $A = [A^1 | \dots | A^n] \in M(m \times n, \mathbb{C})$. Defina:

$$\|A\|_p = \begin{cases} \left(\sum_{i,j} |a_{ij}|^p \right)^{1/p}, & \text{si } 1 \leq p < \infty, \\ \text{máx}_{i,j} |a_{ij}|, & \text{si } p = \infty, \end{cases} \quad \|A\|_p = \begin{cases} \text{máx}_j \|A^j\|_p, & \text{si } 1 \leq p < \infty, \\ \text{máx}_j \|A^j\|_\infty, & \text{si } p = \infty. \end{cases}$$

¿Para qué valores de $p \in [1, \infty]$ es $\|\cdot\|_p$, respectivamente $\|\cdot\|_p$, una norma matricial? No olvide considerar el axioma de submultiplicatividad $\|AB\|_p \leq \|A\|_p \|B\|_p$. Si para ciertos valores de $p, \|\cdot\|_p$, respectivamente $\|\cdot\|_p$, no es una norma matricial, ¿se podría modificar la definición correspondiente para que sí lo fuera?

5. LA NORMA NATURAL DE UNA MATRIZ.

5.1 Toda matriz $A \in M(m \times n, \mathbb{C})$ puede ser considerada como (la representación de) un operador lineal $\mathcal{A} : \mathbb{C}^n \rightarrow \mathbb{C}^m$ (con respecto a sendas bases de \mathbb{C}^n y \mathbb{C}^m). Si consideramos cualquier par de normas en \mathbb{C}^n y \mathbb{C}^m respectivamente, ambas denotadas por $\|\cdot\|$ para mayor simplicidad notacional, entonces la norma de la matriz A en cuanto operador lineal, o la norma matricial inducida por las normas de \mathbb{C}^n y \mathbb{C}^m , o la norma natural de A se define como:

$$\|A\| = \sup_{0 \neq x \in \mathbb{C}^n} \frac{\|A \cdot x\|_m}{\|x\|_n} = \text{máx}_{\|x\|_n=1} \|A \cdot x\|_m.$$

Note que $x \in \mathbb{C}^n$ y $A \cdot x \in \mathbb{C}^m$. Que la norma de A en cuanto operador lineal sea una norma matricial, i.e., que satisfaga los axiomas (i)-(iv) de más arriba, es una consecuencia de las

propiedades generales de las normas vectoriales.

TAREA: *Verifique esta asección.*

5.2 La norma natural $\|A\|$ de A satisface:

$$\|A.x\| \leq \|A\| \|x\|, \quad \text{para todo } x \in \mathbb{C}^n.$$

TAREA: *Verifique esta asección.*

5.3 Una manera de probar que una función dada $\varphi : M(m \times n, \mathbb{C}) \rightarrow \mathbb{R}$ es una norma matricial consiste en demostrar que φ es inducida por una norma vectorial.

TAREA: *Dé algunos ejemplos.*

5.4 La norma natural minimal $\|A\|_{\text{inf}}$ de A se define como:

$$\|A\|_{\text{inf}} = \inf\{\|A\| : \|\cdot\| \text{ es una norma natural}\}.$$

TAREA: *Verifique que $\|A\|_{\text{inf}}$ es una norma matricial.*

5.5 Sea $A \in M(n \times n, \mathbb{C})$. El *spectrum* $\sigma(A)$ de la matriz *cuadrada* A se define como:

$$\sigma(A) = \{\lambda \in \mathbb{C} : \lambda \text{ es valor propio de } A\} = \{\lambda \in \mathbb{C} : \det(A - \lambda I) = 0\}. \quad (11)$$

Como se sabe del Algebra, $\det(A - \lambda I) = 0$ es una ecuación —llamada *ecuación característica* de la matriz A — de grado n y, por lo tanto, tiene n raíces en el plano complejo \mathbb{C} , contando multiplicidades. Por consiguiente, $1 \leq \#\sigma(A) \leq n$. El radio espectral $\rho(A)$ de la matriz A se define como:

$$\rho(A) = \max_{\lambda \in \sigma(A)} |\lambda|. \quad (12)$$

El radio espectral de una matriz A establece una cota inferior para las normas naturales de A .

Lema 1. $\|A\|_{\text{inf}} \geq \rho(A)$.

Demostración. Consideremos \mathbb{C}^n equipado con una norma vectorial $\|\cdot\|$. Sea $A \in M(n \times n, \mathbb{C})$ y $u \in \mathbb{C}^n$ un vector propio normalizado (i.e., $\|u\| = 1$) de A con valor propio $\lambda \in \sigma(A)$, i.e., $A.u = \lambda u$. Entonces se tiene $\|A.u\| = \|\lambda u\| = |\lambda|$. Luego, para la norma matricial natural $\|\cdot\|$ inducida por la norma vectorial $\|\cdot\|$ se tiene

$$\|A\| := \max_{\|x\|=1} \|A.x\| \geq \|A.u\| = |\lambda|, \quad \text{para todo } \lambda \in \sigma(A).$$

De aquí resulta entonces $\|A\| \geq \rho(A)$ y, en consecuencia, $\|A\|_{\text{inf}} \geq \rho(A)$. ■

Lema 2. Para todo $\varepsilon > 0$ existe una norma matricial natural $\|\cdot\|$ tal que $\|A\| \leq \rho(A) + \varepsilon$.

La demostración del Lema 2 requiere de un resultado de Schur que establece que toda matriz cuadrada A es *unitariamente equivalente* a una matriz triangular superior cuyos coeficientes diagonales son los valores propios de A en cualquier orden prescrito.

Teorema 1. Sea $A \in M(n \times n, \mathbb{C})$ una matriz con valores propios $\lambda_1, \dots, \lambda_n$ en cualquier orden prescrito. Entonces existe una matriz unitaria $U \in M(n \times n, \mathbb{C})$ tal que $U^*AU = T = [t_{ij}]_{n \times n}$ es una matriz triangular superior con coeficientes diagonales $t_{ii} = \lambda_i$, $i = 1 : n$. La matriz U no es necesariamente única. Si $A \in M(n \times n, \mathbb{R})$ y si todos los valores propios de A son reales, entonces U se puede elegir real y ortogonal.

Para una demostración del importante Teorema 1 véase [2, §2.3, Theorem 2.3.1, p. 79].

Demostración del Lema 2. [2, §5.6, Lema 5.6.10, p. 297]. Por el teorema de triangularización de Schur 1 existe una matriz unitaria U y una matriz triangular superior $\Delta = [d_{ij}]_{n \times n}$ tal que

$A = U^* \Delta U$. Sea $D_t := \text{diag}[t, t^2, t^3, \dots, t^n]$, $t > 0$. Consideremos ahora la matriz:

$$\begin{aligned}
 D_t \Delta D_t^{-1} &= \text{diag}[t, t^2, t^3, \dots, t^n] \begin{bmatrix} \lambda_1 & d_{12} & d_{13} & \dots & d_{1,n} \\ 0 & \lambda_2 & d_{23} & \dots & d_{2,n} \\ 0 & 0 & \lambda_3 & \dots & d_{3,n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & d_{n-1,n} \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix} \text{diag}[t^{-1}, t^{-2}, t^{-3}, \dots, t^{-n}] \\
 &= \begin{bmatrix} t\lambda_1 & td_{12} & td_{13} & \dots & td_{1,n} \\ 0 & t^2\lambda_2 & t^2d_{23} & \dots & t^2d_{2,n} \\ 0 & 0 & t^3\lambda_3 & \dots & t^3d_{3,n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & t^{n-1}d_{n-1,n} \\ 0 & 0 & 0 & \dots & t^n\lambda_n \end{bmatrix} \text{diag}[t^{-1}, t^{-2}, t^{-3}, \dots, t^{-n}] \\
 &= \begin{bmatrix} \lambda_1 & t^{-1}d_{12} & t^{-2}d_{13} & \dots & t^{-n+1}d_{1,n} \\ 0 & \lambda_2 & t^{-1}d_{23} & \dots & t^{-n+2}d_{2,n} \\ 0 & 0 & \lambda_3 & \dots & t^{-n+3}d_{3,n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & t^{-1}d_{n-1,n} \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix}.
 \end{aligned}$$

Para $t > 0$ suficientemente grande, la suma de los valores absolutos de los términos fuera de la diagonal de $D_t \Delta D_t^{-1}$ es menor que ε , de modo que:

$$\begin{aligned}
 \|D_t \Delta D_t^{-1}\|_1 &= \text{máx} \left\{ |\lambda_1|, |\lambda_2| + t^{-1}|d_{12}|, |\lambda_3| + \sum_{i=1}^2 t^{-3+i}d_{i,3}, \dots, |\lambda_n| + \sum_{i=1}^{n-1} t^{-n+i}d_{i,n} \right\} \\
 &\leq \text{máx} \{ |\lambda_1|, |\lambda_2| + \varepsilon, |\lambda_3| + \varepsilon, \dots, |\lambda_n| + \varepsilon \} \\
 &\leq \text{máx} \{ |\lambda_1|, |\lambda_2|, |\lambda_3|, \dots, |\lambda_n| \} + \varepsilon \\
 &\leq \rho(A) + \varepsilon.
 \end{aligned}$$

Luego, si definimos la norma matricial $\|\cdot\|$ mediante:

$$\|B\| = \|D_t U^* B U D_t^{-1}\|_1, \quad B \in M(n \times n, \mathbb{C}),$$

para $t = t(B) > 0$ suficientemente grande se tiene:

$$\|B\| = \|D_t U^* B U D_t^{-1}\|_1 \leq \rho(U^* B U) + \varepsilon \leq \rho(B) + \varepsilon, \quad B \in M(n \times n, \mathbb{C}).$$

En particular, la matriz $B \in M(n \times n, \mathbb{C})$ satisface esta desigualdad para $t = t(A) > 0$ suficientemente grande. Ello completa la demostración. ■

5.6 Demuestre o refute: Si $n = m$ (i.e., si A es una matriz cuadrada de dimensión n), entonces se tiene $\rho(A) < 1$ si y sólo si $\lim_{p \rightarrow \infty} \|A^p\| = 0$ para alguna norma natural $\|\cdot\|$.

6. La norma espectral de una matriz. La norma espectral $\|A\|_\sigma$ de la matriz cuadrada A se define como:

$$\|A\|_\sigma = \sqrt{\rho(A^* A)} = \text{máx} \left\{ \sqrt{\lambda} : \lambda \text{ es un valor propio de } A^* A \right\}. \quad (13)$$

Nótese que $A^* A$ es una matriz Hermitiana semi-positiva definida. En efecto, $(A^* A)^* = A^* A^{**} = A^* A$ y para todo $x \in \mathbb{C}^n$ se tiene $x^*(A^* A)x = (Ax)^*(Ax) = \|Ax\|_2^2 \geq 0$. Como se sabe del Algebra Lineal, los valores propios de las matrices Hermitianas semi-positivas definidas son reales no-negativos. Luego $\sigma(A^* A) \subset \mathbb{R}^+ \cup \{0\}$, de modo que $\sqrt{\lambda}$ es un número real no-negativo bien definido para todo $\lambda \in \sigma(A^* A)$.

TAREA: Verifique que la norma espectral es una norma matricial i.e., que satisface los axiomas (i)-(iv) de más arriba.

Desarrollo. 1) La verificación se reduce a demostrar que la norma espectral de una matriz A coincide con la norma natural inducida por la norma Euclidea usual (norma l^2) definida tanto en el dominio como en el recorrido de la aplicación lineal $x \mapsto y = Ax$.

2) Sea $B \in M(n \times n, \mathbb{C})$ una matriz dada. Consideremos la función:

$$\varphi(x) = \frac{x^* B x}{x^* x}, \quad 0 \neq x \in \mathbb{C}^n,$$

denominada *cuociente de Rayleigh*. Veamos algunas propiedades de la función $\varphi(x)$. Escribiendo:

$$\varphi(x) = \frac{x^*}{(x^* x)^{1/2}} B \frac{x}{(x^* x)^{1/2}}, \quad 0 \neq x \in \mathbb{C}^n.$$

se observa que $\varphi(x)$ se puede definir, equivalentemente, como:

$$\varphi(x) = x^* B x, \quad x \in \mathbb{C}^n \text{ con } \|x\|_2 = 1,$$

de modo que $\varphi(x)$ está definida sobre la esfera unitaria de \mathbb{C}^n . Si $u \in \mathbb{C}^n$ es un vector propio de B , con valor propio λ , entonces se tiene:

$$\varphi(u) = \frac{u^* B u}{u^* u} = \frac{\lambda u^* u}{u^* u} = \lambda.$$

La función $\varphi(x)$ es acotada:

$$|\varphi(x)| = \frac{|x^* B x|}{x^* x} \leq \frac{\|x\|_2 \|B x\|_2}{\|x\|_2^2} \leq \frac{\|x\|_2 \|B\|_{\mathcal{N}} \|x\|_2}{\|x\|_2^2} = \|B\|_{\mathcal{N}},$$

donde $\|B\|_{\mathcal{N}}$ denota la norma natural de la matriz B con respecto a la norma $\|\cdot\|_2$ tanto en el dominio como en el recorrido. La función $\varphi(x)$ es, además, diferenciable en $\mathbb{C}^n \setminus \{0\}$. La derivada parcial de $\varphi(x)$ con respecto a la j -ésima componente x_j de x viene dada por:

$$\frac{\partial \varphi}{\partial x_j} = \frac{\frac{\partial}{\partial x_j}(x^* B x)}{x^* x} - \frac{(x^* B x) \frac{\partial}{\partial x_j}(x^* x)}{(x^* x)^2} = \frac{2(Bx)_j}{x^* x} - \frac{(x^* B x) 2x_j}{(x^* x)^2} = \frac{2}{x^* x} [Bx - \varphi(x)x]_j,$$

donde $[\cdot]_j$ denota la j -ésima componente. Luego, el gradiente de $\varphi(x)$ viene dado por:

$$\nabla \varphi(x) = \frac{2}{x^* x} [Bx - \varphi(x)x],$$

y por consiguiente:

$$x \text{ es un vector propio de } B \text{ con valor propio } \varphi(x) \Leftrightarrow \nabla \varphi(x) = 0, \tag{14}$$

Luego, los vectores propios $x \in \mathbb{C}^n$ de B son precisamente los puntos estacionarios de la función $\varphi(x)$ y los valores propios de B son los valores de la función $\varphi(x)$ en los puntos estacionarios.

3) Sea $B = A^* A$. Evidentemente B es una matriz Hermitiana semi-positiva definida. Los valores propios de B son reales no-negativos. En efecto, si $u \in \mathbb{C}^n$ es un vector propio de $A^* A$ con valor propio λ , i.e. $(A^* A)u = \lambda u$, entonces se tiene:

$$0 \leq u^*(A^* A)u = u^* \lambda u = \lambda \|u\|_2^2,$$

de donde se desprende que $\lambda \geq 0$. Sobre la esfera unitaria de \mathbb{C}^n se tiene:

$$\varphi(x) = x^*(A^* A)x = (Ax)^*(Ax) = \|Ax\|_2^2 \geq 0, \quad x \in \mathbb{C}^n \text{ con } \|x\|_2 = 1,$$

En particular, si $u \in \mathbb{C}^n$ es un vector propio *unitario* (i.e., $\|u\|_2 = 1$) de $A^* A$, con valor propio λ , se tiene:

$$\varphi(u) = u^*(A^* A)u = u^* \lambda u = \lambda \|u\|_2^2 = \lambda \geq 0,$$

4) Consideremos ahora el máximo de la función $\varphi(x)$ sobre la esfera unitaria de \mathbb{C}^n :

$$\max_{\|x\|_2=1} \varphi(x) = \max_{\|x\|_2=1} \|Ax\|_2^2 \geq 0.$$

Sea $u \in \mathbb{C}^n$ con $\|u\|_2 = 1$ un punto de la esfera unitaria de \mathbb{C}^n donde la función $\varphi(x)$ adopta el máximo (absoluto). Evidentemente, este punto u debe ser un punto estacionario de $\varphi(x)$ y, por lo tanto en vista de (14), u debe ser un vector propio de $A^* A$ con valor propio $\lambda_{\max} := \varphi(u)$.

5) En resumen, de las consideraciones precedentes se desprende que:

$$\begin{aligned} \varphi(u) &= \lambda_{\max} = \max \{ \lambda : \lambda \in \sigma(A^*A) \} = \rho(A^*A) = \|A\|_{\sigma}^2 \\ &\| \\ \max_{\|x\|_2=1} \varphi(x) &= \max_{\|x\|_2=1} \|Ax\|_2^2 = \|A\|_{\mathcal{N}}^2, \end{aligned}$$

donde debe tenerse presente que los valores propios de A^*A son no negativos. Luego, la norma espectral coincide con la norma natural inducida por las normas Euclidianas en el dominio y el recorrido. ■

7. [4, p. 24, ex. 3.3] Las normas de vectores y las p -normas de matrices satisfacen varias desigualdades que a menudo incluyen las dimensiones m o n . En cada una de las sub-preguntas que siguen en este ejercicio, verifique (¡o refute!) la desigualdad anotada y exhiba un ejemplo de matriz o vector no nulos que corroboren su resultado. Si la desigualdad es válida, determine una condición para que se cumpla el caso límite de igualdad y exhiba un ejemplo que verifique esa igualdad. En este ejercicio $x \in \mathbb{K}^m$ y $A \in M(m \times n, \mathbb{K})$.

(a) $\|x\|_{\infty} \leq \|x\|_2$, (b) $\|x\|_2 \leq \sqrt{m}\|x\|_{\infty}$, (c) $\|A\|_{\infty} \leq \sqrt{n}\|A\|_2$, (d) $\|A\|_2 \leq \sqrt{m}\|A\|_{\infty}$.

8. [4, p. 24, ex. 3.4] Sea $A \in M(m \times n, \mathbb{K})$ y sea $B \in M(\mu \times \nu, \mathbb{K})$ con $\mu \leq m$ y $\nu \leq n$.

(a) Demuestre que B puede obtenerse a partir de A mediante multiplicación por matrices apropiadas de “aniquilación” de filas o columnas.

(b) Usando el resultado precedente demuestre que $\|B\|_p \leq \|A\|_p$ para todo $p \in \mathbb{R}$ con $1 \leq p \leq \infty$.

9. [4, p. 24, ex. 3.6] Sea $\|\cdot\|$ una norma cualquiera sobre \mathbb{C}^m . La correspondiente *norma dual* $\|\cdot\|'$ sobre \mathbb{C}^m se define mediante la fórmula $\|x\|' = \sup_{\|y\|=1} |y^*x|$.

(a) Demuestre que $\|\cdot\|'$ es efectivamente una norma vectorial.

(b) Demuestre que, dado un vector $x \in \mathbb{C}^m$, existe un vector no nulo $z \in \mathbb{C}^m$ tal que $|z^*x| = \|z\|'\|x\|$.

(b) Sean $x, y \in \mathbb{C}^m$ con $\|x\| = \|y\| = 1$ dados. Usando (b), demuestre que existe una matriz de tipo *rank*-uno $B = yz^*$ tal que $Bx = y$ y $\|B\| = 1$, donde $\|B\|$ es la norma de la matriz B inducida por la norma vectorial $\|\cdot\|$.

Descomposición de Valor Singular

2. DESCOMPOSICIÓN DE VALOR SINGULAR

1. Considere la matriz $A = \begin{bmatrix} 3 & 2 \\ -1 & 2 \\ 2 & 0 \end{bmatrix}$.

(a) Determine los valores singulares de A .

(b) Determine los vectores singulares *derechos* de A , i.e., los vectores-columna de la matriz \widehat{V} de la SVD *reducida* $\widehat{U}\widehat{\Sigma}\widehat{V}^*$ de A . Note bien que lo pedido son los vectores-columna de \widehat{V} y no los de \widehat{V}^* .

(c) Determine los vectores singulares *izquierdos* de A , i.e., los vectores-columna de la matriz \widehat{U} de la SVD *reducida* $\widehat{U}\widehat{\Sigma}\widehat{V}^*$ de A .

(d) Verifique que la SVD de A que Ud. obtuvo es correcta.

DESARROLLO.

(a) *Preámbulo.* Consideremos la SVD *reducida* de $A \in M(m \times n, \mathbb{C})$:

$$A = U\Sigma V^* \quad \text{y, equivalentemente,} \quad A^* = V\Sigma U^*,$$

donde $U \in M(m \times n, \mathbb{C})$ y $V \in M(n \times n, \mathbb{C})$ tienen columnas respectivamente ortonormales (i.e., $(U^j)^*U^k = \delta_{jk}$ y $(V^j)^*V^k = \delta_{jk}$)¹, y

$$\Sigma = \text{diag} [\sigma_1, \sigma_2, \dots, \sigma_n] \in M(n \times n, \mathbb{C}), \quad \text{con } \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0,$$

es una matriz diagonal. De aquí se desprende que:

$$A^*A = V\Sigma U^*U\Sigma V^* = V\Sigma^2 V^*,$$

donde $A^*A \in M(n \times n, \mathbb{C})$. Luego, $(A^*A).V = V.\Sigma^2$, i.e.:

$$\begin{aligned} (A^*A).[V^1|V^2|\dots|V^n] &= [V^1|V^2|\dots|V^n].\text{diag} [\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2] \\ &= [\sigma_1^2 V^1, \sigma_2^2 V^2, \dots, \sigma_n^2 V^n], \end{aligned}$$

donde V^k es la k -ésima columna de V . Por consiguiente, los vectores-columna de la matriz V de la SVD de A deben satisfacer:

$$(A^*A).V^k = \sigma_k^2 V^k, \quad k = 1 : n,$$

i.e., el k -ésimo vector-columna de la matriz V debe ser un vector propio de la matriz A^*A con valor propio σ_k^2 . Esta observación permite obtener la matriz V con cierta facilidad. En el caso real habrá que considerar, evidentemente, la matriz $A^T A$.

- (b) Cálculo de $A^T A$. Se obtiene directamente $A^T A = \begin{bmatrix} 14 & 4 \\ 4 & 8 \end{bmatrix}$.

Cálculo de los valores propios de $A^T A$ mediante la solución de la ecuación $\det(A^T A - \lambda I) = 0$. Se obtiene $\lambda_1 = \sigma_1^2 = 16$ y $\lambda_2 = \sigma_2^2 = 6$.

Cálculo de los valores singulares de $A^T A$ tomando raíz de los valores propios ya determinados. Se obtiene $\sigma_1 = 4$ y $\sigma_2 = \sqrt{6}$.

- (c) Como se sabe, los vectores singulares *derechos* v_i de A corresponden a los vectores propios de la matriz $A^T A$, normalizados a norma 1. Luego, hay que resolver el sistema $A^T A.v = \lambda v$.

Caso $\lambda = \lambda_1 = 16$: se obtiene $v_1 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$.

Caso $\lambda = \lambda_2 = 6$: se obtiene $v_2 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$.

Nota: Los *signos* de v_1 y v_2 no están unívocamente determinados, pero ello no influye en el desarrollo ulterior.

- (d) Análogamente, los vectores singulares *izquierdos* de A se obtienen resolviendo el sistema $A\hat{V} = \hat{U}\hat{\Sigma}$, donde $\hat{V} = [v_1|v_2]$, $\hat{\Sigma} = \text{diag}[\sigma_1, \sigma_2]$ y los σ_i son los valores singulares obtenidos en (a). Luego, basta resolver el sistema

$$\underbrace{\begin{bmatrix} 8/\sqrt{5} & 1/\sqrt{5} \\ 0 & \sqrt{5} \\ 4/\sqrt{5} & -2/\sqrt{5} \end{bmatrix}}_{=AV} = \begin{bmatrix} 3 & 2 \\ -1 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \\ u_{31} & u_{32} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & \sqrt{6} \end{bmatrix} = \underbrace{\begin{bmatrix} 4u_{11} & \sqrt{6}u_{12} \\ 4u_{21} & \sqrt{6}u_{22} \\ 4u_{31} & \sqrt{6}u_{32} \end{bmatrix}}_{=U\Sigma}$$

para $u_1 = [u_{11}, u_{21}, u_{31}]^T$ y $u_2 = [u_{12}, u_{22}, u_{32}]^T$. De consiguiente:

$$\begin{aligned} u_1 &= \begin{bmatrix} u_{11} \\ u_{21} \\ u_{31} \end{bmatrix} = \frac{1}{\sigma_1} \text{columna}_1(AV) = \frac{1}{4} \text{columna}_1 \left(\begin{bmatrix} 8/\sqrt{5} & 1/\sqrt{5} \\ 0 & \sqrt{5} \\ 4/\sqrt{5} & -2/\sqrt{5} \end{bmatrix} \right) = \begin{bmatrix} 2/\sqrt{5} \\ 0 \\ 1/\sqrt{5} \end{bmatrix} \\ u_2 &= \begin{bmatrix} u_{12} \\ u_{22} \\ u_{32} \end{bmatrix} = \frac{1}{\sigma_2} \text{columna}_2(AV) = \frac{1}{\sqrt{6}} \text{columna}_2 \left(\begin{bmatrix} 8/\sqrt{5} & 1/\sqrt{5} \\ 0 & \sqrt{5} \\ 4/\sqrt{5} & -2/\sqrt{5} \end{bmatrix} \right) = \begin{bmatrix} 1/\sqrt{30} \\ \sqrt{5/6} \\ -\sqrt{2/15} \end{bmatrix} \end{aligned}$$

¹ δ_{jk} es el llamado *símbolo de Kronecker* definido mediante $\delta_{jk} = \begin{cases} 1, & \text{si } j = k, \\ 0, & \text{si } j \neq k. \end{cases}$

(e) Verificación:

$$\widehat{U} \widehat{\Sigma} \widehat{V}^T = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{30} \\ 0 & \sqrt{5/6} \\ 1/\sqrt{5} & -\sqrt{2/5} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & \sqrt{6} \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ -1 & 2 \\ 2 & 0 \end{bmatrix} = A$$

2. (a) Construya su propia demostración para el siguiente teorema o estudie a fondo las demostraciones que proponen algunos de los textos de referencia para este importante resultado del Algebra Lineal Computacional. Asegúrese de haber comprendido correctamente las demostraciones que estudie.

Theorem 1. Sea $A \in M(m \times n, \mathbb{C})$. Entonces existen matrices unitarias $U \in M(m \times m, \mathbb{C})$ y $V \in M(n \times n, \mathbb{C})$, y una matriz diagonal $\Sigma = \text{diag} [\sigma_1, \sigma_2, \dots, \sigma_n] \in M(m \times n, \mathbb{C})$ con $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$,² tales que $A = U\Sigma V^*$.

Los valores singulares están siempre unívocamente determinados.

Si A es una matriz cuadrada y los σ_k son todos distintos, entonces los los vectores singulares izquierdos:

$$U^j = j\text{-ésima columna de la matriz } U, \quad j = 1 : m = n,$$

los vectores singulares derechos:

$$V^k = k\text{-ésima columna de la matriz } V, \quad j = 1 : n = m,$$

están unívocamente determinados excepto por signos complejos (i.e., factores escalares complejos de módulo 1).

(b) Las demostraciones más populares del resultado precedente, son inductivas. Estudie la posibilidad de traducir algunas de esas demostraciones inductivas en algoritmos computacionales que permita obtener la SVD de una matriz dada A .

3. (Cf. Sea $A \in M(m \times n, \mathbb{C})$ una matriz dada y $A = U\Sigma V^*$ su SVD *reducida*. Sea $r = |\{\sigma_j : \sigma_j > 0\}|$. Demuestre o refute, y en todo caso, discuta:

(a) $\text{rank } A = r$.

(b) $\text{Range}(A) = \text{Col}(A) = \langle U^1, \dots, U^r \rangle$, donde $U^j = j$ -ésimo vector-columna de U .

(c) $\text{Null}(A) = \langle V^{r+1}, \dots, V^n \rangle$, donde $V^k = k$ -ésimo vector-columna de V .

(d) $\|A\|_2 = \sigma_1$; $\|A\|_{\text{Frobenius}} = \sqrt{\sigma_1^2 + \dots + \sigma_r^2}$; $\min_{x \neq 0} \frac{\|A \cdot x\|_2}{\|x\|_2} = \sigma_n$ ($m \geq n$).

(e) Los valores singulares σ_j *no nulos* vienen dados por $\sigma_j = \sqrt{\lambda_j}$, donde λ_j es el j -ésimo valor propio *no nulo* de la matriz $A^*A \in M(n \times n, \mathbb{C})$ o de la matriz $AA^* \in M(m \times m, \mathbb{C})$. Ambas matrices tiene los mismos valores propios no nulos. Los valores propios de A^*A y AA^* son reales y no negativos.

(f) Si $A^* = A$, los valores singulares de A coinciden con los valores absolutos de los valores propios de A .

(g) Si $A \in M(m \times m, \mathbb{C})$, $|\det A| = \prod_{j=1}^m \sigma_j$.

4. [4, Lect. 5, p. 35] APROXIMACIÓN DE BAJO RANGO. Sea $A \in M(m \times n, \mathbb{C})$ una matriz dada y $A = U\Sigma V^*$ su SVD *reducida*. Sea $p = \min\{m, n\}$ y $r = |\{\sigma_j : \sigma_j > 0\}|$. Discuta, verifique, corrija o eventualmente refute los siguientes argumentos y resultados:

²Obsérvese que la matriz Σ no es, en general, cuadrada. No obstante, aún en ese caso se acostumbra a decir que Σ es una matriz diagonal. Cuando Σ no es cuadrada, sus filas de por debajo de la fila n , si $m \geq n$, o sus columnas a la derecha de la columna m , si $m \leq n$, son todas nulas.

(a) Dado que $\Sigma = \text{diag} [\sigma_1, \dots, \sigma_r, 0, \dots, \underset{\downarrow}{0}] = \sum_{j=1}^r \text{diag} [0, \dots, 0, \underset{\downarrow}{\sigma_j}, 0, \dots, \underset{\downarrow}{0}]$, se tiene:

$$A = U\Sigma V^* = \sum_{j=1}^r U \cdot \text{diag} [0, \dots, 0, \underset{\downarrow}{\sigma_j}, 0, \dots, \underset{\downarrow}{0}] \cdot V^* = \sum_{j=1}^r \sigma_j U^j \cdot (V^j)^*,$$

(b) Sea $A_\nu = \sum_{j=1}^\nu \sigma_j U^j \cdot (V^j)^*$, $\nu = 0 : r$. Entonces:

$$A - A_\nu = \sum_{j=\nu+1}^r \sigma_j U^j \cdot (V^j)^* = \underbrace{[U^{\nu+1} \dots U^r]}_{U_{\nu+1}} \cdot \underbrace{\text{diag} [\sigma_{\nu+1}, \dots, \sigma_r]}_{\Sigma_{\nu+1}} \cdot \underbrace{\begin{bmatrix} (V^{\nu+1})^* \\ \vdots \\ (V^r)^* \end{bmatrix}}_{V_{\nu+1}^*}$$

constituye una SVD para $A - A_\nu$: $A - A_\nu = U_{\nu+1} \cdot \Sigma_{\nu+1} \cdot V_{\nu+1}^*$. Por definición de la SVD se tiene entonces:

$$\|A - A_\nu\|_2 = \text{primer (mayor) valor singular} = \sigma_{\nu+1}. \quad (15)$$

(c) Sea $E = \langle V^1, \dots, V^{\nu+1} \rangle$ el espacio de \mathbb{C}^n engendrado por los primeros $\nu + 1$ vectores singulares derechos. Para todo $w \in E$ se tiene entonces:

$$\begin{aligned} (A.w)^* \cdot (A.w) &= w^* A^* A w = w^* V \Sigma U^* U \Sigma V^* w = \underbrace{w^* V}_{z^*} \Sigma^2 \underbrace{V^* w}_z = z^* \Sigma^2 z \\ &= \sigma_1^2 |z_1|^2 + \dots + \sigma_{\nu+1}^2 |z_{\nu+1}|^2 \geq \underbrace{(|z_1|^2 + \dots + |z_{\nu+1}|^2)}_{=\|z\|^2 = \|w\|^2} \sigma_{\nu+1}^2. \end{aligned}$$

Nótese que $\|z\|^2 = \|V^* \cdot w\|^2 = \|w\|^2$, pues la matriz $V \in M(n \times n, \mathbb{C})$ es unitaria y, por lo tanto, preserva la norma Euclidea. Por consiguiente se tiene:

$$\|A.w\|^2 \geq \sigma_{\nu+1}^2 \|z\|^2 = \sigma_{\nu+1}^2 \|w\|^2 \quad \forall w \in E. \quad (16)$$

(d) Sobre la base de los resultados precedentes se puede demostrar el siguiente:

Theorem 2. Para todo $\nu = 0 : r$ se tiene:

$$\|A - A_\nu\|_2 \stackrel{(15)}{=} \sigma_{\nu+1} \stackrel{\downarrow}{=} \inf_{\substack{B \in M(m \times n, \mathbb{C}) \\ \text{rank } B \leq \nu}} \|A - B\|_2. \quad (17)$$

Si $\nu = \min\{m, n\}$, hacemos $\sigma_{\nu+1} = 0$.

Demostración. Sólo hay que probar la segunda igualdad en (17). Supongamos que exista una matriz $B \in M(m \times n, \mathbb{C})$ con $\text{rank } B \leq \nu$ tal que:

$$\|A - B\|_2 < \|A - A_\nu\|_2 = \sigma_{\nu+1}.$$

Consideremos entonces el subespacio:

$$W = \text{Null}(B) \subseteq \mathbb{C}^n, \quad \text{con} \quad \dim W = n - \text{rank } B \geq n - \nu.$$

Para $w \in W$ se tiene entonces $B.w = 0$, $A.w = (A - B).w$, y:

$$\|A.w\|_2 = \|(A - B).w\|_2 \leq \|A - B\|_2 \|w\|_2 < \sigma_{\nu+1} \|w\|_2.$$

Luego, W es un subespacio de \mathbb{C}^n de dimensión a lo menos $n - \nu$ donde se cumple $\|A.w\|_2 < \sigma_{\nu+1} \|w\|_2$.

Por otro lado, de (c) sabemos que el subespacio E de \mathbb{C}^n generado por los $\nu + 1$ vectores singulares derechos V^j , $j = 1 : \nu + 1$, de A , posee las siguientes propiedades:

$$\dim E = \nu + 1 \quad \text{y} \quad \|A.w\|_2 \stackrel{(16)}{\geq} \sigma_{\nu+1} \|w\|_2 \quad \forall w \in E.$$

En vista de $\dim W \geq n - \nu$ y $\dim E = \nu + 1$, los subespacios W y E de \mathbb{C}^n deben tener una intersección que no se reduce al subespacio trivial $\{0\}$. En efecto, si $W \cap E \neq \{0\}$, entonces se tendría:

$$n + 1 = (n - \nu) + (\nu + 1) \leq \dim W + \dim E \leq \dim \mathbb{C}^n = n,$$

que obviamente es una contradicción. Existe, por lo tanto, un vector $0 \neq w \in W \cap E$ que debe satisfacer simultáneamente $\|A.w\|_2 < \sigma_{\nu+1} \|w\|_2$, por estar en W , y $\|A.w\|_2 \geq \sigma_{\nu+1} \|w\|_2$, por estar en E . Ello representa, evidentemente, otra contradicción. ■

Nota. El lector diligente seguramente habrá observado que la demostración precedente sólo prueba la desigualdad $\sigma_{\nu+1} \leq \inf_{\substack{B \in M(m \times n, \mathbb{C}) \\ \text{rank } B \leq \nu}} \|A - B\|_2$ en (17). ¿Cómo se obtendría la otra desigualdad?

5. Sea $A \in M(n \times n, R)$ una matriz positiva definida. Para $i_k \in \mathbb{N}$ con $1 \leq i_1 < i_2 < \dots < i_p \leq n$ sea $A[i_1, \dots, i_p]$ la submatriz de A formada por las filas i_1, \dots, i_p y las columnas i_1, \dots, i_p de la matriz A . Note, entonces, que la diagonal de $A[i_1, \dots, i_p]$ está contenida en la diagonal de A . Se dice que $A[i_1, \dots, i_p]$ es una *submatriz principal* de A .

(a) Demuestre o refute, y en cualquier caso, discuta: $\det A[i_1, \dots, i_p] > 0$ para toda selección de los índices $1 \leq i_1 < i_2 < \dots < i_p \leq n$.

(b) ¿Es equivalente (a) con el hecho que A es positiva definida?

(c) ¿Qué relación hay entre los valores singulares de las matrices $A[i_1, \dots, i_p]$ y los valores singulares de la matriz A ?

(d) Sea $V(A, \gamma)$ el volumen del elipsoide $\{x \in \mathbb{R}^n : x^T . A . x \leq \gamma\}$. Calcule el cociente $\frac{V(A, \gamma)}{\det(A)^{1/2}}$.

6. [2, p. 415] Sea $x \in \mathbb{C}^n$ un vector no nulo dado. Considere el vector x como una matriz $X \in M(n \times 1, \mathbb{K})$. Sean $W = [1] \in M(1 \times 1, K)$, $\Sigma = [\|x\|_2, 0, \dots, 0]^T \in M(n \times 1, \mathbb{K})$, $V = \left[\frac{x}{\|x\|_2}, v_2, \dots, v_n \right] \in M(n \times 1, \mathbb{K})$, donde v_2, \dots, v_n es un sistema de $n - 1$ vectores ortonormales cualesquiera, todos ellos ortogonales al vector x . Demuestre que una descomposición de valor singular de la matriz X viene dada por $X = V\Sigma W^*$.

7. [2, p. 416] Si $A \in M(n \times n, \mathbb{K})$ es no singular, demuestre que el siguiente procedimiento produce una descomposición de valor singular $A = V\Sigma W^*$:

(i) Forme la matriz Hermitiana positiva definida AA^* y determine una diagonalización unitaria $AA^* = U\Lambda U^*$ hallando los valores propios positivos $\lambda_i, i = 1, \dots, n$, y el correspondiente sistema de $u_i, i = 1, \dots, n$, de vectores propios ortonormales.

(ii) Defina $\Sigma := \Lambda^{1/2}$ y $V = U = [u_1, \dots, u_n]$.

(iii) Defina $W = A^*V\Sigma^{-1}$.

a) Demuestre todo. En particular, demuestre que W es unitaria con $A = V\Sigma W^*$.

b) Escriba un programa que permita obtener la descomposición de valor singular.

c) Usando su programa, obtenga descomposiciones de valor singular para tres o cuatro matrices de dimensiones 3, 5, y 8, por ejemplo. Para resolver esta parte del ejercicio el estudiante probablemente querrá aplicar el algoritmo descrito en el ejercicio precedente. Converse con sus colegas que abordaron ese ejercicio para que le faciliten los códigos correspondientes.

d) Usando un programa *ad hoc* verifique que el producto de las matrices de las factorizaciones que obtenga, efectivamente dan las matrices originales.

8. [2, p. 416] Sea $A \in M(n \times n, K)$ una matriz no necesariamente no-singular. Demuestre que el siguiente procedimiento produce una descomposición de valor singular $A = V\Sigma W^*$:

(i) Sea $I \in M(n \times n, K)$ la matriz identidad. Entonces existe una constante $c = c(A)$ tal que la matriz $A_\varepsilon = A + \varepsilon I$ es no-singular para todo $\varepsilon \in]0, c[\subset \mathbb{R}$. *Sugerencia:* Observe las raíces de la ecuación característica $\det(A - \lambda I) = 0$ en el plano complejo.

- (ii) Elija ahora $\varepsilon \in]0, c[\subset \mathbb{R}$. Aplique el procedimiento del ejercicio anterior para obtener una descomposición de valor singular $A_\varepsilon = V_\varepsilon \Sigma_\varepsilon W_\varepsilon^*$
- (iii) Aplique ahora el siguiente resultado, *que no se pide demostrar*:

Lema Auxiliar. Sea $\{U_p\}_{p \in \mathbb{N}} = \left[u_{ij}^p \right]_{p \in \mathbb{N}}$ una sucesión de matrices unitarias en $M(n \times n, \mathbb{K})$. Entonces existe una subsucesión $\{U_{p_k}\}_{k \in \mathbb{N}} = \left[u_{ij}^{p_k} \right]_{k \in \mathbb{N}}$ de la sucesión original, tal que

$$\lim_{k \rightarrow \infty} u_{ij}^{p_k} = u_{ij}^0, \quad \forall i, j = 1, \dots, n,$$

donde la matriz límite $U_0 = \left[u_{ij}^0 \right] \in M(n \times n, \mathbb{K})$ también es unitaria.

Sobre la base de este lema auxiliar se obtiene una sucesión $\{\varepsilon_k\}_{k \in \mathbb{N}}$ en $]0, c[$, una matriz diagonal $\Sigma \in M(n \times n, \mathbb{K})$ con coeficientes no negativos, y matrices unitarias $V, W \in M(n \times n, \mathbb{K})$ tales que:

$$\lim_{k \rightarrow \infty} \varepsilon_k = 0, \quad \lim_{k \rightarrow \infty} V_{\varepsilon_k} = V, \quad \lim_{k \rightarrow \infty} \Sigma_{\varepsilon_k} = \Sigma, \quad \lim_{k \rightarrow \infty} W_{\varepsilon_k} = W.$$

- (iv) Demuestre que $A = V \Sigma W^*$.

9. [2, p. 421, pr. 7] Sea $A \in M(m \times n, \mathbb{K})$ una matriz dada con descomposición de valor singular $A = V \Sigma W^*$. Defina $A^\dagger = W \Sigma^\dagger V^*$, donde Σ^\dagger es la transpuesta de Σ en la cual los valores singulares positivos de A se reemplazan por sus valores recíprocos. Demuestre que:

- (a) AA^\dagger y $A^\dagger A$ son hermitianas, (b) $AA^\dagger A = A$, (c) $A^\dagger AA^\dagger = A^\dagger$.

Demuestre que $A^\dagger = A^{-1}$ si A es cuadrada y no-singular. La matriz A^\dagger se llama *inversa generalizada* o *inversa de Moore-Penrose* de la matriz A . Se puede demostrar que la inversa generalizada existe para toda matriz A , incluso para matrices cuadradas singulares o para matrices no cuadradas.

Demuestre, además, que A^\dagger queda unívocamente determinada por las propiedades (a), (b), (c).

10. [2, p. 422, pr. 10] La descomposición de valor singular se puede obtener sin utilizar explícitamente vectores ni valores propios. Los vectores singulares (de izquierda y de derecha) y los valores singulares pueden construirse directamente a partir de la llamada *caracterización variacional de la norma espectral* (que es, simplemente, la definición usual de la norma de un operador lineal): si $A \in M(n \times n, \mathbb{K})$ es una matriz dada, entonces se define:

$$\|A\|_2 = \sup_{\|x\|_2=1} \|Ax\|_2. \quad \text{(Norma Espectral)}$$

- (a) Sea $n \geq 2$ y suponga que $B \in M(n \times n, \mathbb{K})$ tiene la forma:

$$B = \begin{bmatrix} \sigma_1 & w^* \\ 0 & X \end{bmatrix}, \quad \text{con } \sigma_1 = \|B\|_2, \quad w \in \mathbb{C}^{n-1}, \quad X \in M((n-1) \times (n-1), \mathbb{K}).$$

Demuestre que $w = 0$. *Sugerencia:* Si $\sigma > 0$, considere $\zeta = \frac{1}{(\sigma_1^2 + w^*w)^{1/2}} \begin{bmatrix} \sigma_1 \\ w \end{bmatrix}$, demuestre que $\|B\zeta\|_2^2 \geq \sigma_1^2 + w^*w$, y use la fórmula (Norma Espectral). Además, ¡consulte [4]!

(b) Sea $A \in M(n \times n, \mathbb{K})$, defina $\sigma_1 = \|A\|_2$, y use la fórmula (Norma Espectral) para probar que existe un vector unitario $\|x\|_1$ tal que $\|Ax_1\|_2 = \sigma_1$. Defina $y_1 = \sigma_1^{-1} Ax_1$ para la parte (c). *Nota:* Posiblemente el estudiante quiera usar aquí el teorema de Weierstrass para funciones continuas definidas sobre conjuntos compactos.

(c) Sean $V_1, W_1 \in M(n \times n, \mathbb{K})$ matrices unitarias cuyas primeras columnas son x_1 e y_1 respectivamente. Demuestre que $V_1^* A W_1$ tiene norma espectral σ_1 y que tiene la forma de la matriz B de la parte (a). Concluya que $V_1^* A W_1 = \begin{bmatrix} \sigma_1 & 0 \\ 0 & X \end{bmatrix}$.

(d) Formule un proceso de “deflación inductiva” de A introduciendo columnas y filas con “ceros fuera de la diagonal” (cuando se las observa en la matriz A) mediante pre- y post-multiplicación por

matrices unitarias apropiadas. Verifique que el proceso de inducción termina tras un número finito de pasos y que al final se obtiene la descomposición de valor singular de A .

(e) ¿Qué ocurre si la matriz original A no es cuadrada?

11. (Cf. [2, Theor. 7.3.2, p. 412]) Considere una descomposición de valor singular $A = U\Sigma V^*$ de una matriz cuadrada $A \in M(n \times n, \mathbb{K})$. Escriba:

$$A = U\Sigma V^* = (U\Sigma U^*) (UV^*) = PW, \quad \text{con } P = U\Sigma U^* \quad \text{y} \quad W = UV^*. \quad (18)$$

Demuestre o refute, y en todo caso, discuta:

- (a) P es positiva semidefinida, i.e., $x^*Px \geq 0$ (note que dice “ \geq ” y no “ $>$ ”) para todo $x \in \mathbb{K}^n$, con $x \neq 0$.
- (b) $\text{rank } P = \text{rank } A$.
- (c) W tiene filas ortonormales, i.e. $WW^* = I$.
- (d) P está unívocamente determinada y satisface $P = (AA^*)^{1/2}$. Dé una interpretación razonable para esta última fórmula.
Suponga ahora que la matriz original A no es cuadrada, digamos $A \in M(m \times n, \mathbb{K})$ con $m \geq n$. A “bruta fuerza”, haga cuadradas (de $m \times m$) todas las matrices (exceptuando V^* que conviene dejarla como una matriz de $m \times n$) de la descomposición de valor singular, agregando 0’s a diestra y siniestra.
- (e) Examine si la factorización (18) y las propiedades (a)-(d) siguen siendo válidas o no.
- (f) La nueva matriz W ¿está unívocamente determinado cuando $\text{rank } A = n$?
- (g) Si A es real, ¿es posible tomar P y W como matrices reales?
- (h) ¿Qué pasa si $m \leq n$?

Como el lector sospechará, (casi) todas las respuestas deberían ser afirmativas. La factorización obtenida se conoce como “representación polar” de una matriz dada.

(i) Escriba un programa para obtener la representación polar de cualquier matriz.

12. [2, p. 415][4, p. 37, exc. 5.2] Usando SVD³ demuestre que toda matriz $A \in M(m \times n, \mathbb{C})$ es el límite (¿cuál norma sería apropiada para calcular el límite?) de una sucesión de matrices $A_k = \left(a_{ij}^{(k)} \right)_{m \times n} \in M(m \times n, \mathbb{C})$, $k \in \mathbb{N}$, de rango completo. En otras palabras, demuestre que el conjunto de las matrices de rango completo es denso en $M(m \times n, \mathbb{C})$.

13. [4, p. 37, exc. 5.3] [For all students!!!] Considere la matriz $\begin{bmatrix} -2 & 11 \\ -10 & 5 \end{bmatrix}$.

- (a) Determine s.c.P.L.P.⁴ una SVD real de A de la forma $A = U\Sigma V^T$. La SVD no es única (compruebe este hecho). Así, encuentre la SVD que tiene el menor número de signos negativos en U y V .
- (b) Haga una tabla con los valores singulares, los vectores singulares izquierdos, y los vectores singulares derechos de A . Dibuje cuidadosamente la bola unitaria de \mathbb{R}^2 y su imagen bajo A , junto con los vectores singulares, indicando claramente las coordenadas de sus vértices. Use sistemas de coordenadas claramente rotulados con los nombres de las variables involucradas y todos los signos convencionales del Dibujo Técnico que aseguren una buena comprensión de sus diagramas.
- (c) Determine las normas 1, 2, ∞ , y de Frobenius-Hilbert-Schmidt de la matriz A .
- (d) Determine A^{-1} indirectamente vía SVD.
- (e) Verifique que: $\det A = \lambda_1 \lambda_2$ y $|\det A| = \sigma_1 \sigma_2$.
- (f) Determine el área del elipsoide que resulta como imagen de la bola unitaria de \mathbb{R}^2 bajo A .

14. [4, p. 37, exc. 5.4] Suponga que $A \in M(m \times m, \mathbb{C})$ tiene una SVD $A = U\Sigma V^*$. Encuentre una EVD⁵ $A = X\Lambda X^{-1}$ de la matriz hermitiana $\begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix} \in M(2m \times 2m, \mathbb{C})$.

³ “singular value decomposition”

⁴ s.c.P.L.P. = sólo con papel, lápiz y pensamiento

⁵ “eigenvalue decomposition”

15. Cf.: <http://www.computer.org/cise/homework/v5n3/index.htm>

James G. Nagy and Dianne P. O'Leary, [3]. Los llamados *problemas inversos* constituyen uno de los desafíos más exigentes en la computación aplicada a la ciencia y a la ingeniería, debido a que involucran la determinación de los parámetros de un sistema observado sólo indirectamente. Un ejemplo de tales problemas es, por ejemplo, la determinación de un sistema a partir solamente de su *espectro*. Otro ejemplo típico es la determinación de posibles fracturas en un (es)tanque exclusivamente a partir de mediciones de tipo sonar.

El problema que se quiere discutir aquí es el siguiente: dada una imagen borrosa y un modelo lineal para la borrosidad, reconstruir la imagen original. Este problema inverso lineal ilustra el impacto del *mal condicionamiento* en la selección de algoritmos. La herramienta principal para este proyecto es la SVD.

(1) MAL CONDICIONAMIENTO. Consideremos un sistema de ecuaciones lineales:

$$K \cdot f = g, \quad K \in M(n \times n, \mathbb{R}), \quad f, g \in \mathbb{R}^n.$$

Supondremos que la matriz K está escalada de modo que su mayor valor singular es $\sigma_1 = 1$. Si el menor valor singular σ_n es aproximadamente 0, entonces K está mal condicionada. Distinguiremos dos tipos de mal condicionamiento:

- La matriz K se considera *numéricamente de rango deficiente* si existe un j tal que $\sigma_j \gg \sigma_{j+1} \approx \dots \approx \sigma_n \approx 0$, i.e., si hay una brecha ostensible entre los valores singulares grandes y los pequeños.
- Si los valores singulares decaen a 0 sin ninguna brecha especial en el espectro, se dice que el sistema lineal $K \cdot f = g$ es un *problema mal condicionado discreto*

El cálculo de soluciones aproximadas muy cercanas a las soluciones exactas en un problema mal condicionado discreto del tipo $K \cdot f = g$ es, por lo general, un problema muy difícil, especialmente debido a que en muchas aplicaciones reales, g no se conoce exactamente. En realidad la data típicamente tiene la forma:

$$g = K \cdot f + \eta,$$

donde η representa un ruido (desconocido) o errores de medición. El objetivo es, entonces, dada una matriz mal condicionada K y un vector g , calcular una (buena) aproximación del vector desconocido f .

La solución ingenua de $K \cdot f = g$, usualmente, no es apropiada cuando K es muy mal condicionada, como en el caso que nos interesa. En estos casos se usa un procedimiento de *regularización* para hacer el problema menos sensible al ruido. En algunos círculos se habla de *robustificación* del proceso.

(2) REGULARIZACIÓN DE TIKHONOV. El método de regularización más conocido es la llamada *regularización de Tikhonov* que calcula una solución del problema de mínimos cuadrados amortiguado:

$$\min_f \left\{ \|g - Kf\|_2^2 + \alpha^2 \|f\|_2^2 \right\}. \tag{19}$$

El término $\alpha^2 \|f\|_2^2$ impone una penalización al tamaño de la solución. Dicha penalización reduce el efecto de los valores singulares pequeños.

(3) PROBLEMA 1. Demuestre que la ecuación (19) es equivalente al problema de mínimos cuadrados lineal:

$$\min_f \left\| \begin{bmatrix} g \\ 0 \end{bmatrix} - \begin{bmatrix} K \\ \alpha I \end{bmatrix} f \right\|_2^2. \tag{20}$$

El escalar α (llamado *parámetro de regularización*) controla el grado de *suavidad* de la solución. Evidentemente, $\alpha = 0$ implica ausencia de regularización y en este caso la solución de (20) estará muy probablemente corrupta con mucho ruido. Por otro lado, si α es grande, la solución calculada no podrá ser una buena aproximación de la solución exacta f . La elección de un valor apropiado para α no es algo trivial. Varios algoritmos han sido propuestos en la literatura pero aquí se usará un método manual.

- (4) PROBLEMA 2. Demuestre que si K tiene una SVD $K = U\Sigma V^T$, entonces la ecuación (20) puede transformarse en el problema de cuadrados mínimos equivalente:

$$\min_f \left\| \begin{bmatrix} \hat{g} \\ 0 \end{bmatrix} - \begin{bmatrix} \Sigma \\ \alpha I \end{bmatrix} \hat{f} \right\|_2^2, \quad \text{con } \hat{f} = V^T f, \quad \hat{g} = U^T g. \quad (21)$$

- (5) PROBLEMA 3. Obtenga una fórmula para la solución de la ecuación (21). *Hint*: haga cero la derivada de la función de minimización y resuelva para f .
 Todo esto constituye un algoritmo para determinar la solución de Tikhonov de un problema mal condicionado discreto.
- (6) SVD TRUNCADA. Otro método para regularizar el problema consiste en truncar la SVD. El siguiente Problema 4 ilustra cómo expresar la solución al problema de mínimos cuadrados en términos de la SVD.
- (7) PROBLEMA 4. Demuestre que la solución del problema:

$$\min_f \left\{ \|g - Kf\|_2^2 \right\}. \quad (22)$$

viene dada por:

$$f_{ls} = V\Sigma^\dagger U^T \equiv \sum_{i=1}^n \frac{u_i^T g}{\sigma_i} v_i, \quad (23)$$

donde u_i es la i -ésima columna de U , y v_i es la i -ésima columna de V .

Se puede ver que hay problemas en f_{ls} si un pequeño valor de σ_i divide un término $u_i^T g$ que contiene errores. En tales casos, f_{ls} quedará dominado por errores.

Para eludir esta dificultad, Richard Hanson y James Varah han sugerido truncar la expansión (23):

$$f_{ls} = \sum_{i=1}^p \frac{u_i^T g}{\sigma_i} v_i, \quad (24)$$

para algún valor $p < n$.

Ahora tenemos todos los ingredientes para resolver el problema de eliminación de la borrosidad en procesamiento de imágenes. Supongamos que tenemos una imagen borrosa o ruidosa G junto con algún conocimiento del operador de borrosidad, y queremos construir la imagen verdadera original F . Este es un ejemplo de un problema mal condicionado discreto en el cual los vectores en el sistema lineal $g = K.f + \eta$ representan los arreglos bidimensionales (F y G) de las imágenes apilados por columnas para formar vectores (por ejemplo, debajo de la primera columna de F viene la segunda columna de F y así sucesivamente hasta concluir con la n -ésima columna, obteniéndose de esta manera un vector de n^2 componentes). En la notación de Matlab se tendría:

$$f = \text{reshape}(F, n, 1), \quad g = \text{reshape}(G, n, 1). \quad (25)$$

El objetivo de este problema es, dados K y G , reconstruir una aproximación de la imagen desconocida F .

Si suponemos que F y G contienen $\sqrt{n} \times \sqrt{n}$ pixeles, entonces f y g son vectores de largo n y K es una matriz de $n \times n$ que representa la borrosidad en la imagen. Por lo general esta matriz es muy grande para usar SVD. Sin embargo, en algunos casos se puede escribir K como un producto de Kronecker $K = A \otimes B$ y entonces se puede usar SVD.

- (8) DOS PALABRITAS ACERCA DE LOS PRODUCTOS DE KRONECKER. El producto de Kronecker $A \otimes B$, donde A y B son matrices de $m \times m$ se define mediante:

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1m}B \\ a_{21}B & a_{22}B & \dots & a_{2m}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mm}B \end{bmatrix}. \quad (26)$$

Theorem 3. Si $A = U_A \Sigma_A V_A^T$ y $B = U_B \Sigma_B V_B^T$, entonces $K = U \Sigma V^T$ donde $U = U_A \otimes U_B$, $\Sigma = \Sigma_A \otimes \Sigma_B$ y $V = V_A \otimes V_B^T$.

Tarea: Demuestre el teorema precedente.

Por consiguiente, si una matriz grande es el producto de Kronecker de dos matrices más pequeñas, entonces se puede calcular la SVD de la matriz grande. En la página “web” de J.G. Nagy y D.P. O’Leary mencionada al comienzo de este ejercicio se puede hallar un programa Matlab de muestra que ilustra la propiedad discutida.

Para resolver el problema de eliminación de la borrosidad de una imagen hay que operar cuidadosamente con matrices más pequeñas pues, de otro modo, rápidamente aparecen problemas de memoria o almacenamiento de data.

- (9) **PROBLEMA 5.** Escriba un programa que toma matrices A , B y una imagen G , y calcule aproximaciones de la imagen F mediante la regularización de Tikhonov y la SVD truncada. Para cada uno de estos algoritmos, realice algunos experimentos para hallar el valor del parámetro de regularización (α para Tikhonov y p para la SVD truncada) que da la imagen más nítida. La referida página “web” de J.G. Nagy y D.P. O’Leary contiene datos de muestra: una imagen borrosa G y matrices A , B . La tarea consiste en restaurar la imagen hasta el punto que se pueda leer el texto que contiene. Compare la efectividad de los dos algoritmos.

3. MIS DUDAS CASTELLANAS

“Las lenguas cambian de continuo, y lo hacen de modo especial en su componente léxico. Por ello los diccionarios nunca están terminados: son una obra viva que se esfuerza en reflejar la evolución registrando nuevas formas y atendiendo a las mutaciones de significado.” (<http://www.rae.es/>)

alcuza. (Del ár. hisp. *alkúza*, este del ár. clás. *kūzah*, este del arameo *kūz[ā]*, y este del persa *kuze*).
1. f. Vasija de barro, de hojalata o de otros materiales, generalmente de forma cónica, en que se guarda el aceite para diversos usos. **2.** f. *Am.* **vinagreras** (|| pieza con dos frascos para aceite y vinagre).

glacial. (Del lat. *glaciālis*). **1.** adj. Helado, muy frío. **2.** adj. Que hace helar o helarse. **3.** adj. Frío, desafecto, desabrido. **4.** adj. Geogr. Dicho de la tierra o del mar: Que está en las zonas glaciales.

glaciar. (Del fr. *glacier*). **1.** m. Masa de hielo acumulada en las zonas de las cordilleras por encima del límite de las nieves perpetuas y cuya parte inferior se desliza muy lentamente, como si fuese un río de hielo. **2.** adj. Perteneciente o relativo al glaciar.

LECCION 6^a. *Acentuación de este, ese, aquel...*

(<http://www.aulafacil.com/OrtografEspañola/Lecc-6-ortog.htm>)

En Castellano existen los siguientes demostrativos:

Singular	Plural
Este, esta, esto	Estos, estas
Ese, esa, eso	Esos, esas
Aquel, aquella, aquello	Aquellos, aquellas

Si estas palabras sustituyen a un nombre funcionan como *pronombres demostrativos*. Pero si acompañan a un nombre, entonces funcionan como *adjetivos demostrativos*:

Adjetivo demostrativo (<i>no se acentúan</i>)	Pronombre demostrativo (<i>(se acentúan)</i>)
<i>Ese</i> coche es el mío	<i>Ése</i> es mi coche
¿Quién es <i>aquel</i> individuo?	¿Quién es <i>aquél</i> ?
<i>Aquella</i> niña es tu novia	<i>Aquella</i> es tu novia
¿De dónde viene <i>este</i> señor?	¿De dónde viene <i>éste</i> ?
<i>Esta</i> casa es muy bonita	<i>Eso</i> no me lo creo
<i>Este</i> libro es muy caro.	Hay muchos libros, pero <i>éste</i> es el más caro.
<i>Ese</i> colegio es el mejor de la ciudad.	De todos los colegios, <i>ése</i> es el mejor.
<i>Aquel</i> perro me ha mordido.	¿Qué perro te ha mordido? <i>Aquél</i> ha sido.

El pronombre demostrativo se acentúa precisamente para distinguirlo del adjetivo demostrativo. Por eso mismo, el pronombre demostrativo neutro (*esto, eso, aquello*) no se acentúa ya que no hay una forma adjetiva equivalente con la que se pudiera confundir.

Más ejemplos de frases con pronombre demostrativo: *Éste* es el que más me gusta. *Esto* es lo que más me gusta. *Aquél* es el mejor. *Aquello* es impresionante.

Instantes, Noviembre 2006. Marcelo se inclina y susurra al oído de Luis: “*Se dice ‘en relación a, en relación a...*””. Luis, sorprendido, sólo puede musitar: “*No. Se dice ‘con relación a’. Lo he investigado.*” Marcelo, algo molesto, replica: “*Qué porfiado eres. Le preguntaremos a Norman para salir de dudas.*” <http://www.rae.es/> resuelve la duda:

relación. (Del lat. *relatĭo, -ōnis*). **1.** f. Exposición que se hace de un hecho. **2.** f. Conexión, correspondencia de algo con otra cosa. **3.** f. Conexión, correspondencia, trato, comunicación de alguien con otra persona. U. m. en pl. Relaciones de parentesco, de amistad, amorosas, comerciales. **4.** f. Trato de carácter amoroso. U. m. en pl. Tienen relaciones desde hace tiempo. **5.** f. Lista de nombres o elementos de cualquier clase. **6.** f. Informe que generalmente se hace por escrito, y se presenta ante una autoridad. **7.** f. En el poema dramático, trozo largo que dice un personaje, ya para contar o narrar algo, ya con cualquier otro fin. **8.** f. Gram. Conexión o enlace entre dos términos de una misma oración; p. ej., en la frase ‘*amor de madre*’ hay una relación gramatical cuyos dos términos son las voces amor y madre. **9.** f. Mat. Resultado de comparar dos cantidades expresadas en números. **10.** f. Arg. y Ur. En diversos bailes tradicionales, copla que se dicen los integrantes de las parejas. **11.** f. pl. Conocidos o amigos influyentes. Sin relaciones no se puede triunfar en esa profesión.

~ **de ciego.** **1.** f. romance de ciego. **2.** f. coloq. La frívola e impertinente. **3.** f. coloq. Aquello que se recita o lee con monotonía y sin darle el sentido que corresponde.

~ **jurada.** **1.** f. Razón o cuenta que con juramento expreso en ella se da a quien tiene autoridad para exigirla.

relaciones públicas. **1.** f. pl. Actividad profesional cuyo fin es, mediante gestiones personales o con el empleo de las técnicas de difusión y comunicación, informar sobre personas, empresas, instituciones, etc., tratando de prestigiarlas y de captar voluntades a su favor. **2.** f. pl. u. c. sing. com. Persona que desempeña esta profesión.

con ~ a. **1.** loc. prepos. Que tiene conexión o correspondencia con algo. **2.** loc. prepos. **con respecto a.**

en ~ con. **1.** loc. adv. con relación a.

hacer ~ a algo. **1.** fr. Tener con ello conexión aquello de que se trata. **2.** fr. Der. En los pleitos y causas, dar cuenta al tribunal relatando lo esencial de todo el proceso.

catequesis. (Del lat. *catechēsis*, y este del gr. *κατήχησις*). **1.** f. Ejercicio de instruir en cosas pertenecientes a la religión. **2. f. Arte de instruir por medio de preguntas y respuestas.** **3.** f. Lugar o reunión donde se imparte la doctrina cristiana.

catequismo. (Del lat. catechismus, y este del gr. *κατηχισμός*, de *κατηχεῖν*, instruir). **1. m.** catequesis (|| ejercicio de instruir en cosas pertenecientes a la religión). **2. m. catequesis** (|| **arte de instruir por medio de preguntas y respuestas**). **3. m.** desus. catecismo (|| libro de instrucción elemental que contiene la doctrina cristiana). **4. m. ant. catecismo** (|| **obra que contiene la exposición sucinta de alguna ciencia o arte**).

Conviene usar siempre las palabras *κατήχησις*, *κατηχισμός* y *κατηχεῖν* en su sentido griego primigenio de propagar o hacer eco de las enseñanzas ... de la Filosofía Natural, especialmente de la Computación Científica, tan popular en el siglo de oro de la aquella comarca llamada Castalia en honor y recuerdo de la hija de de Aqueloo (!).

Vorläufiges Ende.

REFERENCIAS

- [1] G.H. Golub and C. Van Loan. *Matrix Computations*. John Hopkins University Press, Baltimore-London, 1996 (Third Edition). This encyclopedic work is the must in the field!
- [2] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1994.
- [3] J.G. Nagy and D.P. O'Leary. Image deblurring: I can see clearly now. *Computing in Science and Engineering*, 5(3):82–84, May-June 2003. Copublished by the IEEE Computer Society and the American Institute of Physics.
- [4] L.N. Trefethen and D. Bau III. *Numerical Linear Algebra*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1997. A very good book indeed! More advanced than Strang's book.

LSC/lsc, Valparaíso, 3 de abril de 2008