Lecturas de Optimización Numérica:

PROYECTORES Y PSEUDOINVERSAS DE MOORE-PENROSE

Prof. Dr. Fredy Vides

Scientific Computing Innovation Center, UNAH &
Centre for Analysis of Data-Driven Systems
E-mail: fredy.vides@unah.edu.hn

ÍNDICE

OBJETIVOS

1. Estudiar algunas propiedades fundamentales de los proyectores.
2. Estudiar algunas propiedades fundamentales de las pseudoinversas de Moore-Penrose.

1. PROYECTORES

Definición 1.1. Una matriz $P \in \mathbb{C}^{n\times n}$ se denomina un **proyector** si $P^2 = P = P^*$.

Definición 1.2. Dada una matriz $X \in \mathbb{C}^{n\times m}$, se denomina el rango de X como el número entero determinado por el máximo número de columnas linealmente independientes de X, se denota por $\mathrm{rk}(X)$ el rango de X.

Definición 1.3. Dada una matriz $A \in \mathbb{C}^{n\times m}$, se define el **espacio imagen** de A como el conjunto $\{y = Ax : x \in \mathbb{C}^m\} \subseteq \mathbb{C}^n$.

Dada una matriz $F \in \mathrm{C}^{n\times m}$ cuyas columnas son linealmente idependientes, y dado $x \in \mathbb{C}^m$. El vector Fx es un elemento del espacio generado por la columnas de F. El vector Fx es una proyección ortogonal de un vector $y \in \mathbb{C}^n$ si se cumple la siguiente restricción.

$$F^*(y - Fx) = 0 \tag{1.1}$$

Por (1.1) se tiene que Fx es la proyección de un vector $z \in \mathbb{C}^n$ en el espacio imagen de F siempre que se cumple la siguiente restricción.

$$x = (F^*F)^{-1}F^*z \tag{1.2}$$

Observación 1.4. Es importante observar que, en efecto, $F^*F \in \mathbb{GL}_m$ como consecuencia de las propiedades de la descomposición QR de F, dado que las columnas de F son linealmente independientes.

Ejercicio para el lector 1. Verificar la observación 1.4.

Con base en la observación 1.4, si se define la matriz

$$P_F = F(F^*F)^{-1}F^* \tag{1.3}$$

se pueden realizar las siguientes observaciones.

Observación 1.5. Es posible observar que la matriz P_F definida por (1.3) cumple las siguientes restricciones.

$$\begin{aligned} P_F^2 &= P_F P_F \\ &= F(F^*F)^{-1}F^*F(F^*F)^{-1}F^* \\ &= F(F^*F)^{-1}F^* = P_F \end{aligned} \tag{1.4}$$

$$\begin{aligned} P_F^* &= (F(F^*F)^{-1}F^*)^* \\ &= F((F^*F)^{-1})^*F^* \\ &= F((F^*F)^*)^{-1}F^* \\ &= F(F^*F)^{-1}F^* = P_F \end{aligned} \tag{1.5}$$

Además, por (1.4) y (1.5) es posible observar que la matriz $I - P_F$ cumple las siguientes restricciones.

$$\begin{aligned} (I - P_F)^2 &= (I - P_F)(I - P_F) \\ &= I - P_F - P_F + P_F^2 \\ &= I - 2P_F + P_F = I - P_F \end{aligned} \tag{1.6}$$

$$(I - P_F)^* = I^* - P_F^* = I - P_F \tag{1.7}$$

$$(I - P_F)P_F = P_F - P_F^2 = P_F - P_F = \mathbf{0} \tag{1.8}$$

$$P_F(I - P_F) = P_F - P_F^2 = P_F - P_F = \mathbf{0} \tag{1.9}$$

Observación 1.6. Por la observación 1.5, si P_F es la matriz definida por la ecuación (1.3) para una matriz $F \in \mathbb{C}^{m\times n}$ cuyas columnas son linealmente independientes, entonces P_F y $I - P_F$ son proyectores, y además $P_F(I - P_F) = (I - P_F)P_F = \mathbf{0}$.

Definición 1.7. Dada una matriz $A \in \mathbb{C}^{n\times m}$ cuyas columnas son linealmente independientes, se denotará por P_A el proyector definido por la expresión $P_A = A(A^*A)^{-1}A^*$.

Definición 1.8. Dada una matriz $A = [a_{jk}] \in \mathbb{C}^{n\times n}$ se denomina traza de A el número $\operatorname{tr}(A)$ definido por la siguiente expresión.

$$\operatorname{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$$

Ejercicio para el lector 2. Dado un proyector $P \in \mathbb{C}^{n\times n}$. Demostrar que P cumple con las siguientes condiciones.

- $\lambda(P) \subset \{0, 1\}$
- $\mathrm{tr}(P) = \mathrm{rk}(P)$
- $I - 2P \in \mathbb{U}(n)$
- $\lambda(I - 2P) \subset \{-1, 1\}$

2. Pseudoinversas de Moore-Penrose

Dada una matriz arbitraria A tal que $\mathrm{rk}(A) = r > 0$, si se considera una representación

$$A = FR^*$$

donde tanto F como R tienen r columnas linealmente independientes. Es válido definir las matrices.

$$F_A = P_F = F(F^*F)^{-1}F^*$$
$$F_{A^*} = P_R = R(R^*R)^{-1}R^*$$

Ahora es posible considerar la matriz definida por la siguiente expresión.

$$A^+ = R(R^*R)^{-1}(F^*F)^{-1}F^* \tag{2.1}$$

Observación 2.1. Por cómputo directo es posible verificar lo siguiente.

$$\begin{aligned} AA^+ &= FR^*R(R^*R)^{-1}(F^*F)^{-1}F^* \\ &= F(F^*F)^{-1}F^* = P_F = F_A \end{aligned}$$

También es posible verificar lo siguiente.

$$\begin{aligned} A^+A &= R(R^*R)^{-1}(F^*F)^{-1}F^*FR^* \\ &= R(R^*R)^{-1}R^* = P_R = F_{A^*} \end{aligned}$$

Además, A y A^+ cumplen las siguientes condiciones.

$$AA^+A = A, \quad A^+AA^+ = A^+ \tag{2.2}$$

Definición 2.2. Dada una matriz A tal que $\mathrm{rk}(A) = r > 0$, se denomina **pseudoniversa** (de Moore-Penrose) o **inversa generalizada** de A, la matriz A^+ definida por la expresión (2.1).

Observación 2.3. La pseudo inversa A^+ de una matriz A tal que $\mathrm{rk}(A) = r > 0$ cumple las restricciones (2.2).

2.1. Pseudoinversas y problemas de mínimos cuadrados.

Dada una matriz $A \in \mathbb{C}^{m\times n}$ tal que $\mathrm{rk}(A) = r > 0$, es posible demostrar que A^+ es única. Dado un vector $y \in \mathbb{C}^m$, formalmente A^+ resuelve el siguiente problema de mínimos cuadrados.

$$x = \mathrm{argmin}_{z\in\mathbb{C}^n} \|y - Az\|_2^2$$

Donde tal como se establece en el planteamiento del problema, la solución (minimizador) x es el vector $x \in \mathbb{C}^n$ que permite obtener el mínimo valor posible para la expresión $\|y - Ax\|_2^2$. Dado que el vector Ax es la proyección ortogonal de y sobre el espacio imagen de A. Por tanto la solución requerida es

$$x = A^+y,$$

y además se cumplen las siguientes condiciones.

$$Ax = AA^+y = P_F y$$
$$A^*Ax = A^*AA^+y = A^*P_F y = RF^*F(F^*F)^{-1}F^*y = RF^*y = (FR^*)^*y = A^*y$$

Ejercicio para el lector 3. Dada una matriz $A \in \mathbb{C}^{n\times m}$ tal que $\mathrm{rk}(A) = r > 0$, y dado $y \in \mathbb{C}^n$. Demostrar que existe un proyector $Q \in \mathbb{C}^{n\times n}$ que cumple la siguiente restricción.

$$\|Qy\|_2^2 = \min_{x\in\mathbb{C}^m} \|Ax - y\|_2^2$$

Es posible utilizar el comando `pinv` de Octave para calcular la pseudoinversa (de Moore-Penrose) de una matriz.

Ejemplo 1. Es posible generar una matriz al azar en $\mathbb{C}^{7\times 7}$ de rango 3 utilizando la siguiente secuencia de comandos.

```
>> A=ceil(randn(7,3));
>> A=A*ceil(randn(7,3))'
A =

   0   0   1  -1   0   0   1
   0   0   1  -1   0   0   1
   1   1   2   0   0  -1  -2
  -1   2   2   1   1  -2  -2
  -1   2   3   0   1  -2  -1
  -3   3   4   0   2  -3   0
  -3   3   1   3   2  -3  -3
```

Es posible calcular el rango $\mathrm{rk}(A)$ utilizando la siguiente secuencia de comandos.

```
>> rank(A)
ans =  3
```

Para calcular A^+ es posible utilizar el comando `pinv` como se muestra en la siguiente secuencia de comandos.

```
>> Ap=pinv(A);
```

Podemos verificar **aproximadamente** algunas de las propiedades fundamentales de A^+ utilizando octave, la razón por la que en general la verificación es aproximada se debe, como se ha discutido anteriormente en el curso, a los efectos de la aritmética finita.

Para calcular el valor $\|AA^+A - A\|_2$, es posible escribir la siguiente secuencia de comandos:

```
>> norm(A*Ap*A-A)
ans =    3.8665e-15
```

El proyector P_F determinado por la observación 2.1 puede calcularse aproximadamente con Octave utilizando la siguiente secuencia de comandos basada en la observación 2.1.

```
>> Pf=A*Ap;
```

Es posible verificar aproximadamente algunas de las propiedades fundamentales del proyector P_F utilizando la siguiente secuencia de comandos.

```
>> norm(Pf'-Pf)
ans =    7.7442e-16
```

Esta secuencia de comandos aproxima el valor $\|P_F^* - P_F\|_2$.

```
>> norm(Pf^2-Pf)
ans =    6.5079e-16
```

Esta secuencia de comandos aproxima el valor $\|P_F^2 - P_F\|_2$.

```
>> trace (Pf)
ans =  3
>> rank(Pf)
ans =  3
```

Estas secuencias de comandos calculan los valores $\mathrm{tr}(P_F)$ y $\mathrm{rk}(P_F)$, respectivamente. Es posible observar que tal como lo predice la teoría en el ejercicio para el lector 2,

$$\mathrm{rk}(P_F) = 3 = \mathrm{tr}(P_F).$$

Esta es una de las muchas formas de **invariantes topológicos** que aparecen en álgebra lineal y optimización numérica.

Ejercicio para el lector 4. Dada una matriz $A \in \mathbb{C}^{n\times m}$ tal que $\mathrm{rk}(A) = r > 0$, y dado un vector $y \in \mathbb{C}^n$. Desarrollar un programa Octave que calcula (aproximadamente) el proyector $Q \in \mathbb{C}^{n\times n}$ que cumple la siguiente restricción.

$$\|Qy\|_2^2 = \min_{x\in\mathbb{C}^m} \|Ax - y\|_2^2$$

Ejercicio para el lector 5. Desarrollar un programa Octave que genera un par de matrices de prueba $A \in \mathbb{C}^{n\times m}, y \in \mathbb{C}^n$, que cumplen las condiciones requeridas por el ejercicio para el lector 4.

Referencias

[1] A. S. Householder (1964). The Theory of Matrices in Numerical Analysis. Dover Publications, Inc.
[2] R. L. Burden, D. J. Faires, A. M. Burden. (2017). Análisis Numérico. 10a Ed. Cengage Learning Editores.
[3] Golub, G. H., Van Loan C. F (1996). Matrix Computations (3aEd.). The Johns Hopkins University Press.
[4] Quarteroni A., Saleri F., Gervasio P. (2014). Scientific computing with MATLAB and Octave (Textbook).
[5] D. G. Luenberger, Y. Ye. (2016). Linear and Nonlinear Programming. 4a Ed. Springer International Publishing Switzerland.