EE578B - Convex Optimization - Winter 2021

Homework 2 - Solution

<u>Due Date</u>: Sunday, Jan 24^{th}, 2021 at 11:59 pm

1. **Matrix Rank**

   The column rank of a matrix is the number of linearly independent columns. The row rank of a matrix is the number of linearly independent row.

   (a) **(PTS: 0-2)** Show that the row rank is less than or equal to the column rank.
   Solution: For a matrix $A \in R^{m\times n}$, if the column rank of the matrix is c, we can choose a matrix $C \in R^{m\times c}$ whose columns span the column space of A and write A as

   $$A = CV$$

   where $V \in R^{c\times n}$. The jth column of V are the coordinates of the jth column of A with respect to the columns of C. Note also that we can think of the ith row of C as the coordinates of the ith row of A with respect to the rows of V. Since each row of A is a linear combination of the c rows of V, the row rank is less than or equal to c.

   (b) **(PTS: 0-2)** Show that the col rank is less than or equal to the row rank.
   Solution:
   For a matrix $A \in R^{m\times n}$, if the row rank of the matrix is r, we can choose a matrix $B \in R^{r\times n}$ whose rows span the row space of A and write A as

   $$A = WB$$

   where $W \in R^{m\times r}$. The ith row of W is the coordinates of the ith row of A with respect to the rows of B. Note also that we can think of the jth column of B as the coordinates of the jth column of A with respect to the columns of W. Since each column of A is a linear combination of the r columns of W, the column rank is less than or equal to r.

2. **Grammian Rank**

   (PTS: 0-2) Show that $\text{rank}(A) = \text{rank}(A^T) = \text{rank}(A^TA) = \text{rank}(AA^T)$

   Solution:

   Since row rank = col rank, $\text{rank}(A) = \text{rank}(A^T)$. By the rank-nullity theorem, we know that for $B \in \mathbb{R}^{m\times n}$.

   $$\text{rank}(B) + \dim(\mathcal{N}(B) = n$$

   It follows that if two matrices with the same domain have the same nullspace then their rank is the same. We can show that A and A^TA have the same nullspace. First, if $x \in \mathcal{N}(A)$ then $A^TAx = (A^T)0 = 0$ and $x \in \mathcal{N}(A^TA)$. Secondly, if $x \in \mathcal{N}(A^TA)$ then

   $$A^TAx = 0 \quad \Rightarrow \quad x^TA^TAx = 0 \quad \Rightarrow \quad |Ax|_2^2 = 0$$

The only vector with a 0 2-norm is the 0 vector and thus $Ax = 0$ and $x \in \mathcal{N}(A)$. The same argument equates the rank of A^T and AA^T by replacing A with A^T.

3. **Basis for Domain from Nullspace of A and Range of A^T**

   Consider $A \in \mathbb{R}^{m\times n}$ with $m < n$ and full row rank and a matrix $N \in \mathbb{R}^{n\times n-m}$ with full column rank whose columns span the nullspace of A. Suppose we write a vector $x \in \mathbb{R}^n$ as a linear combination of the rows of A and the columns of N, ie.

$$x = \begin{bmatrix} A^T & N \end{bmatrix} \begin{bmatrix} x_1' \\ x_2' \end{bmatrix}.$$

   for $x_1' \in \mathbb{R}^m$ and $x_2' \in \mathbb{R}^{n-m}$

   (a) **(PTS: 0-2)** Symbollically compute $\begin{bmatrix} A^T & N \end{bmatrix}^{-1}$.

   Hint: Start by checking if $\begin{bmatrix} A^T & N \end{bmatrix}^{-1} = \begin{bmatrix} A^T & N \end{bmatrix}^T$...

   Solution:

   Columns of N span the nullspace of A, so $AN = 0$. A has full row rank, thus, AA^T is invertible. Similarly, N has full column rank, so N^TN is invertible.

$$\begin{aligned}
\begin{bmatrix} A^T & N \end{bmatrix}^T \begin{bmatrix} A^T & N \end{bmatrix} &= \begin{bmatrix} A \\ N^T \end{bmatrix} \begin{bmatrix} A^T & N \end{bmatrix} \\
&= \begin{bmatrix} AA^T & AN \\ (AN)^T & N^TN \end{bmatrix} \\
&= \begin{bmatrix} AA^T & 0 \\ 0 & N^TN \end{bmatrix}
\end{aligned}$$

   Thus,

$$\begin{aligned}
\begin{bmatrix} A^T & N \end{bmatrix}^{-1} &= \begin{bmatrix} AA^T & 0 \\ 0 & N^TN \end{bmatrix}^{-1} \begin{bmatrix} A \\ N^T \end{bmatrix} \\
&= \begin{bmatrix} \left(AA^T\right)^{-1} & 0 \\ 0 & \left(N^TN\right)^{-1} \end{bmatrix} \begin{bmatrix} A \\ N^T \end{bmatrix} \\
&= \begin{bmatrix} \left(AA^T\right)^{-1} A \\ \left(N^TN\right)^{-1} N^T \end{bmatrix}
\end{aligned}$$

   (b) **(PTS: 0-2)** Solve for x_1' and x_2' given A,N, and x.

   Solution:

   According to (a), we have

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} A^T & N \end{bmatrix}^{-1} x = \begin{bmatrix} \left(AA^T\right)^{-1} A \\ \left(N^TN\right)^{-1} N^T \end{bmatrix} x$$

   Hence, $x_1' = \left(AA^T\right)^{-1} Ax, x_2' = \left(N^TN\right)^{-1} N^T x$

4. **Range and Nullspace**

   Let $\mathcal{R}(A)$ and $\mathcal{N}(A)$ represent the range and nullspace of A (and similarly let $\mathcal{R}(A^T)$ and $\mathcal{N}(A^T)$ be the range and nullspace of A^T).

(a) **(PTS: 0-2)** Suppose $y \in \mathcal{R}(A)$ and $x \in \mathcal{N}(A^T)$. Show that $x \perp y$, ie. $x^T y = 0$.

Solution: If $y \in \mathcal{R}(A)$, then there exists z such that $y = Az$. It follows that $x^T y = y^T x = z^T A^T x = z^T 0 = 0$ since $x \in \mathcal{N}(A^T)$. Colloquially, one could summarize this fact by saying that $\mathcal{N}(A^T)$ is orthogonal to the columns of A.

(b) **(PTS: 0-2)** Consider $A \in \mathbb{R}^{5\times 10}$. Suppose A has only 3 linearly independent columns (the other 7 are linearly dependent on the first 3). What is the dimension of $\mathcal{R}(A)$? What is the dimension of $\mathcal{N}(A^T)$ What is the dimension of $\mathcal{N}(A)$? What is the dimension of $\mathcal{R}(A^T)$? (You can state your answers without proof.)

Solution:

$\dim(\mathcal{R}(A)) = \dim(\mathcal{R}(A^T)) = \text{rank}(A) = 3$
$\dim(\mathcal{N}(A)) = 10 - 3 = 7,$
$\dim(\mathcal{N}(A^T)) = 5 - 3 = 2.$

5. **Fundamental Theorem of Linear Algebra Pictures**

For each of the following matrices draw a picture of the domain (either $\mathbb{R}^2$ or $\mathbb{R}^3$) labeling $\mathcal{R}(A^T)$ and $\mathcal{N}(A)$ and a picture of the co-domain (either $\mathbb{R}^2$ or $\mathbb{R}^3$) labeling the $\mathcal{R}(A)$ and $\mathcal{N}(A^T)$.

(PTS: 0-2) $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ **(PTS: 0-2)** $A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$

(PTS: 0-2) $A = \begin{bmatrix} -1 & 1 \\ 1 & 1 \\ 2 & 2 \end{bmatrix}$ **(PTS: 0-2)** $A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \end{bmatrix}$

Solution:

(a) The domain and co-domain are $\mathbb{R}^2$, see Figure. 1 and Figure. 2.

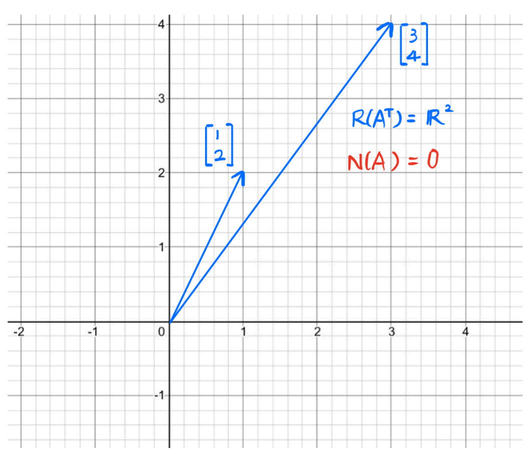

Figure 1: Domain

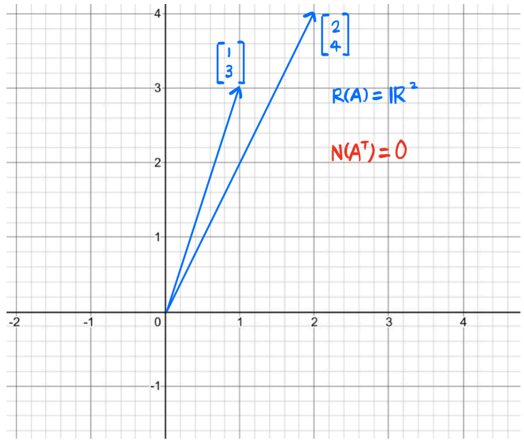

Figure 2: Co-domain

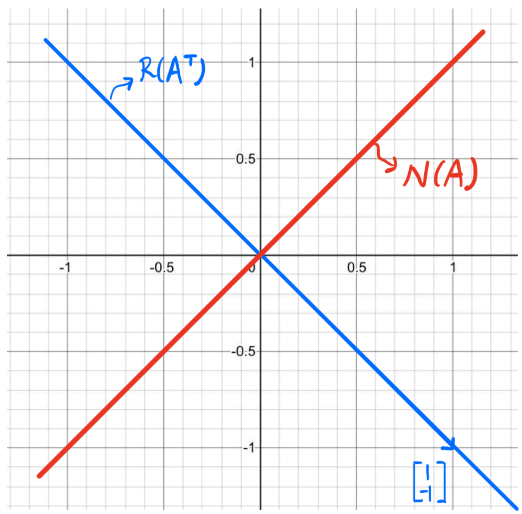

Figure 3: Domain

(b) The domain and co-domain are $\mathbb{R}^2$, see Figure. 3 and Figure. 4.

(c) The domain is $\mathbb{R}^2$ and the co-domain is $\mathbb{R}^3$.

- Co-domain: $R(A)$ is the span of columns of A, i.e.,

$$R(A) = \text{span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \right\}.$$

  $N(A^T)$ is

$$N(A^T) = \text{span}\left\{ \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix} \right\}$$

  Note that in this case, $N(A^T)$ is the normal vector of the plane $R(A)$.

- Domain: $R(A^T)$ is the span of columns of A^T, i.e.,

$$R(A^T) = \text{span}\left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} = \mathbb{R}^2.$$

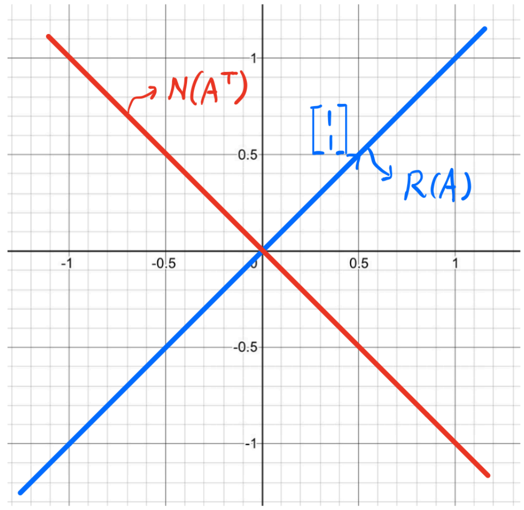

Figure 4: Co-domain

So $N(A) = 0$.

See Figure. 5 and Figure. 6.

(d) The domain is $\mathbb{R}^3$ and the co-domain is $\mathbb{R}^2$.

- Co-domain: $R(A)$ is the span of columns of A, i.e.,

$$R(A) = \text{span}\left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}.$$

$N(A^T)$ is

$$N(A^T) = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

- Domain: We have

$$N(A) = \text{span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\},$$

and

$$R(A^T) = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

See Figure. 7 and Figure. 8.

Note that $R(A^T)$ in this plot is the normal vector of the 2-D plane $N(A)$.

6. **Representations of Affine Sets**

Consider two representations of the same affine set.

$$\textbf{Representation 1:} \qquad \mathcal{X} = \left\{ x \in \mathbb{R}^n \;\middle|\; Ax = b, \right\}$$
$$\textbf{Representation 2:} \qquad \mathcal{X} = \left\{ x \in \mathbb{R}^n \;\middle|\; x = Nz + d,\; z \in \mathbb{R}^{n-m} \right\}$$

where

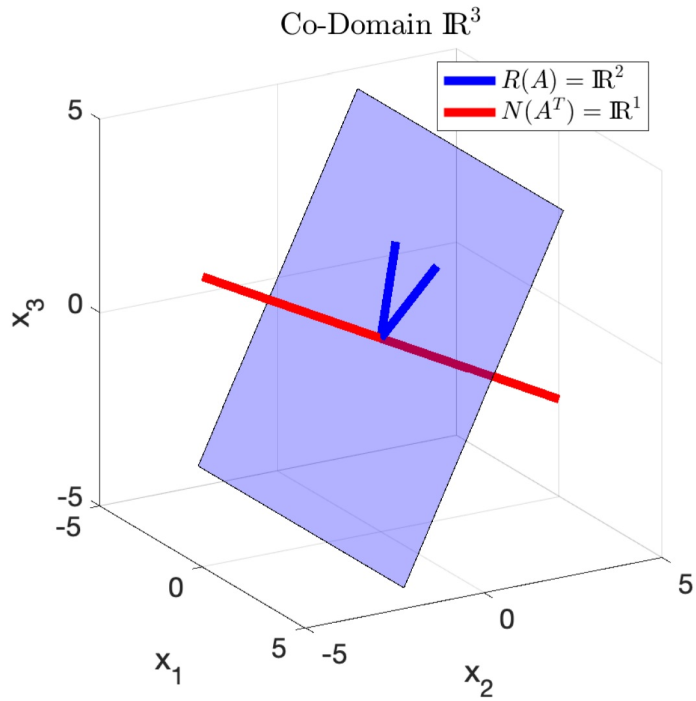

Figure 5: Co-domain (Plot credit to Andrews)

- $A \in \mathbb{R}^{m\times n}$ is fat ($m < n$) and full row rank ($\text{rk}(A) = m$) and $b \in \mathbb{R}^m$
- $N \in \mathbb{R}^{n\times(n-m)}$ is tall, $\mathcal{R}(N) = \mathcal{N}(A)$, and $d \in \mathbb{R}^n$

(a) For each $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, compute $N \in \mathbb{R}^{n\times(n-m)}$ and $d \in \mathbb{R}^n$.
Note: There are many possible N's and d's that work.

Solution:

The columns of N should provide a basis for the nullspace of A. In order to compute a nullspace, we use the technique from Homework 1, Problem 6. If k is the rank of A, we first divide $A \in \mathbb{R}^{m\times n}$ into pieces $A = [A_1\ A_2]$ with $A_1 \in \mathbb{R}^{m\times k}$ and $A_2 \in \mathbb{R}^{m\times(n-k)}$. Note that for each of these matrices $k = m$.

$$N = \begin{bmatrix} A_1^{-1}A_2 \\ -I \end{bmatrix}$$

d is any specific solution to $Ax = b$. One option is the x with the smallest norm such that $Ax = b$. This minimum norm solution is given by $d = A^T(AA^T)^{-1}b$. Another option is to compute the coordinates of b with respect to the basis given by the columns of A_1.

$$d = \begin{bmatrix} A_1^{-1}b \\ 0 \end{bmatrix}$$

Note: Again there are many other N's and d's that would work. The columns of N only need to span the nullspace of A and d needs to solve the equation $Ad = b$.

(PTS: 0-2) $\quad A = \begin{bmatrix} 1 & -2 & 0 \end{bmatrix}, \quad b = 1,$

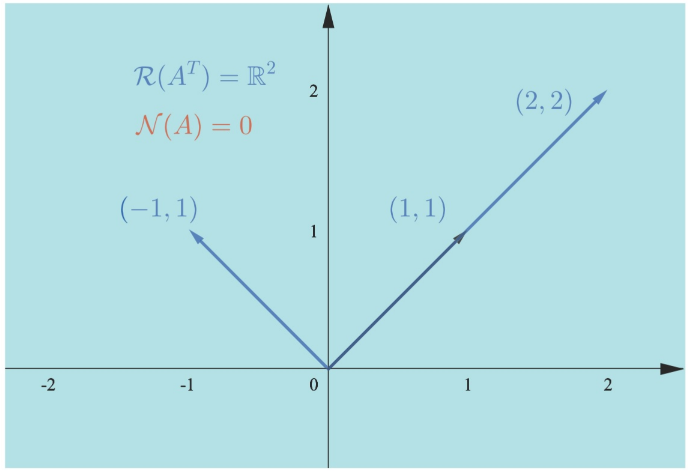

Figure 6: Domain (Plot credit to Weiqian.)

Solution: $A_1 = 1$, $A_2 = [-2\ 0]$.

$$N = \begin{bmatrix} -2 & 0 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad d = A^T(AA^T)^{-1}b = \frac{1}{5}\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \quad \text{OR} \quad d = \begin{bmatrix} A_1^{-1}b \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

(PTS: 0-2) $$A = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

Solution:

$$A_1 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}, \quad A_1^{-1} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$N = \begin{bmatrix} A_1^{-1}A_2 \\ -I \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \qquad d = A^T(AA^T)^{-1}b = \frac{1}{6}\begin{bmatrix} 4 \\ -1 \\ 7 \end{bmatrix}, \quad \text{OR} \quad d = \begin{bmatrix} A_1^{-1}b \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$$

(PTS: 0-2) $$A = \begin{bmatrix} 1 & -2 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Solution:

$$A_1 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}, \quad A_1^{-1} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

$$N = \begin{bmatrix} A_1^{-1}A_2 \\ -I \end{bmatrix} = \begin{bmatrix} 2 & -1 & 3 \\ 1 & -1 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \qquad d = A^T(AA^T)^{-1}b = \begin{bmatrix} -0.083 \\ 0.375 \\ 0.208 \\ -0.292 \\ 0.125 \end{bmatrix} \quad \text{OR} \quad d = \begin{bmatrix} A_1^{-1}b \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

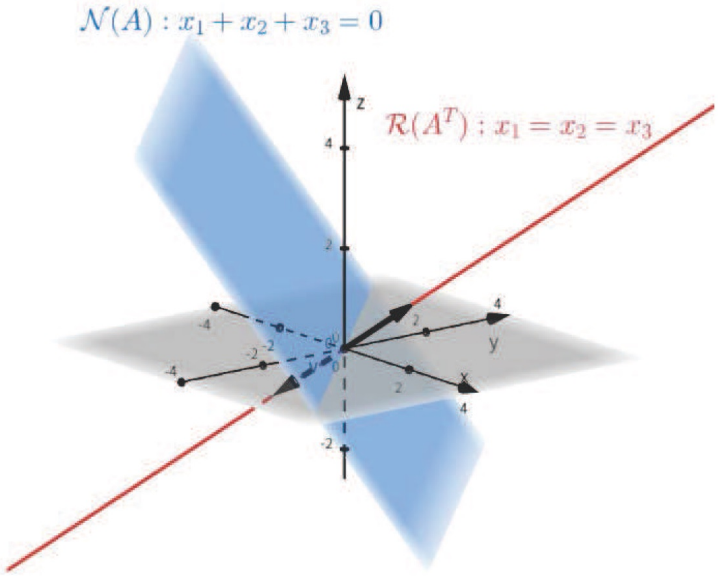

Figure 7: Domain (Plot credit to Weiqian.)

(b) For each $N \in \mathbb{R}^{n\times(n-m)}$ and $d \in \mathbb{R}^n$, compute $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$.
Note 1: There are many possible A's and b's that work.
Note 2: Note that $N^T A^T = 0$, ie. the rows of A should form a basis for the nullspace of N^T.

Solution:

The rows of A should span the space orthogonal to the columns of N. Thus we can compute A^T as a basis for the nullspace of N^T. Let $N^T = [N_1\ N_2]$ with $N_1 \in \mathbb{R}^{(n-m)\times(n-m)}$ and $N_2 \in \mathbb{R}^{(n-m)\times m}$.

$$A^T = \begin{bmatrix} N_1^{-1}N_2 \\ -I \end{bmatrix}$$

Once A is calculated, b is simply given by $b = Ad$

$$\textbf{(PTS: 0-2)} \qquad N = \begin{bmatrix} 1 \\ -2 \\ -2 \end{bmatrix}, \quad d = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$A = \begin{bmatrix} -2 & -1 & 0 \\ -2 & 0 & -1 \end{bmatrix}, \qquad b = \begin{bmatrix} -3 \\ -3 \end{bmatrix}$$

$$\textbf{(PTS: 0-2)} \qquad N = \begin{bmatrix} 1 & 0 \\ -2 & 0 \\ 1 & 1 \end{bmatrix}, \quad d = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

For this example, since the first two columns of N^T aren't linearly independent, we can use the last two columns as a basis.

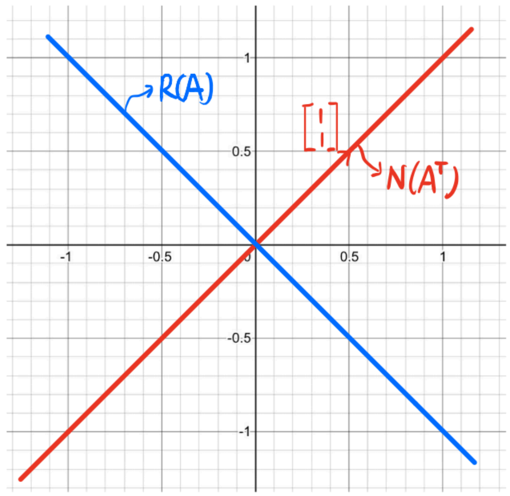

Figure 8: Co-domain

$$N^T = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} N_2 & N_1 \end{bmatrix} \quad \Rightarrow \quad A^T = \begin{bmatrix} -I \\ N_1^{-1} N_2 \end{bmatrix} = \begin{bmatrix} -1 \\ -\frac{1}{2} \\ 0 \end{bmatrix}$$

$$A = \begin{bmatrix} -1 & -\frac{1}{2} & 0 \end{bmatrix}, \qquad b = -\frac{1}{2}$$

(PTS: 0-2) $\quad N = \begin{bmatrix} 1 & -2 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad d = \begin{bmatrix} -1 \\ 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}$

Solution:

$$A = \begin{bmatrix} 1 & 2 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 1 & 3 & 0 & 0 & -1 \end{bmatrix}, \qquad b = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

7. **Equivalent representations of spaces**

- **(PTS: 0-2)** For $A \in \mathbb{R}^{m \times n}$ and $U \in \mathbb{R}^{m \times m}$, invertible, show that $\mathcal{N}(A) = \mathcal{N}(UA)$. OPTIONAL: Comment on how the rows of A relate to the rows of UA.

  Solution: To show that two nullspaces are equivalent we show that an element in one is an element in the other and vice versa. First, if $x \in \mathcal{N}(A)$, then $UAx = U(0) = 0$. Second, suppose $x \in \mathcal{N}(UA)$, $UAx = 0$. Since U is invertible, we have that $Ax = U^{-1}0 = 0$. Note that if U was not invertible then it could be that $Ax \neq 0 \in \mathcal{N}(U)$ and the nullspaces would not be equal. The rows of UA are linear combinations of the rows of A. The rows of U are the coordinates of the rows of UA with respect to the rows of A.

- **(PTS: 0-2)** For $A \in \mathbb{R}^{m\times n}$ and $V \in \mathbb{R}^{n\times n}$, invertible, show that $\mathcal{R}(A) = \mathcal{R}(AV)$.
  OPTIONAL: Comment on how the columns of A relate to the columns of AV.

  Solution: If $y \in \mathcal{R}(A)$, then there exists x such that $y = Ax$. Choosing $x' = V^{-1}x$ shows that $y = Ax = AVV^{-1}x = AVx'$ shows there exists x' such that $y = AVx'$ and thus $y \in \mathcal{R}(AV)$. Similarly, if $y \in \mathcal{R}(AV)$, then there exists x' such that $y = AVx'$. Choosing $x = Vx'$ shows that $y = AVx' = Ax$ shows there exists x such that $y = Ax$ and thus $y \in \mathcal{R}(A)$ The columns of AV are linear combinations of A. The columns of V are the coordinates of the columns of AV with respect to the columns of A.

8. **Vector Derivatives**

Let $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$. Compute $\frac{\partial f}{\partial x}$ for the following functions:

- **(PTS: 0-2)**

$$f(x) = x_1^4 + 3x_1x_2^2 + e^{x_2} + \tfrac{1}{x_1x_2}$$

  Solution:

$$\frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 4x_1^3 + 3x_2^2 + \frac{-1}{x_1^2x_2} & 6_1x_2 + e^{x_2} + \frac{-1}{x_1x_2^2} \end{bmatrix}$$

- **(PTS: 0-2)**

$$f(x) = \begin{bmatrix} \beta x_1 + \alpha x_2 \\ \beta(x_1 + x_2) \\ \alpha^2 x_1 + \beta x_2 \\ \beta x_1 + \frac{1}{\alpha} x_2 \end{bmatrix}$$

  for $\alpha, \beta \in \mathbb{R}$

  Solution:

$$\frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} \\ \frac{\partial f_4}{\partial x_1} & \frac{\partial f_4}{\partial x_2} \end{bmatrix} = \begin{bmatrix} \beta & \alpha \\ \beta & \beta \\ \alpha^2 & \beta \\ \beta & \frac{1}{\alpha} \end{bmatrix}$$

- **(PTS: 0-2)**

$$f(x) = \begin{bmatrix} e^{x^TQx} \\ (x^TQx)^{-1} \end{bmatrix}$$

  for some $Q = Q^T \in \mathbb{R}^{2\times 2}$.

  Solution: Applying the chain rule:

$$\frac{\partial f}{\partial x} = \begin{bmatrix} e^{x^TQx} x^T(Q+Q^T) \\ \frac{-1}{(x^TQx)^2} x^T(Q+Q^T) \end{bmatrix} = \begin{bmatrix} e^{x^TQx} \\ \frac{-1}{(x^TQx)^2} \end{bmatrix} x^T(Q+Q^T)$$