Polytope Geometry

Dan Calderone

***Abstract*—This tutorial paper discusses linear inequality and polytope visualizations in terms of matrix column and row geometry. Basic visualizations and relationships to affine/half-spaces are discussed in both contexts. Particular focus is then given to slack variable representations**

I. Introduction

Visualizing matrix geometry is at the heart of developing spatial intuition for linear algebra and sets the stage for visualization of many vector related topics in modern engineering such as optimization and machine learning.

A matrix is a block of numbers used to represent a linear transformation between vector spaces. Definition of a matrix immediately defines both ”columns” and ”rows” of a matrix which have distinct interpretations relative to the geometry of the linear transformation. In this paper, we seek to show how the geometry of the columns relates to the structure of the linear map. The spatial intuition we will develop will have countless applications in the theory of linear equations, optimization, and other fields.

In the first part of the paper introduces matrix column geometry; each column of a matrix defines where the standard basis vectors (and thus the axes) in the domain map to in the co-domain. Basic examples and intuition are developed for matrix-vector multiplication and also matrix-matrix multiplication.

The bulk of the paper is then divided into two sections: one visualizing co-domain sets and one visualizing domain sets. In the co-domain section, images of basic domain sets are discussed along with the range and adjoint nullspace. Particular focus is then given to image representations of affine spaces. In the domain section, basic pre-images are discussed and then both the nullspace and adjoint range are each given lengthy treatment. Finally, pre-image representation of affine spaces are discussed at length.

We note that this paper assumes familiarity with the notation and vector visualization techniques presented in the following monograph.

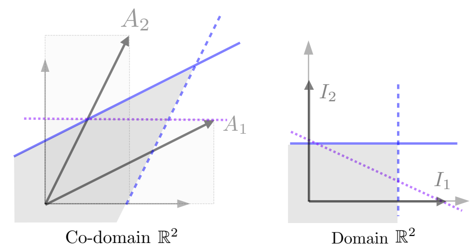
- Vector visualizations

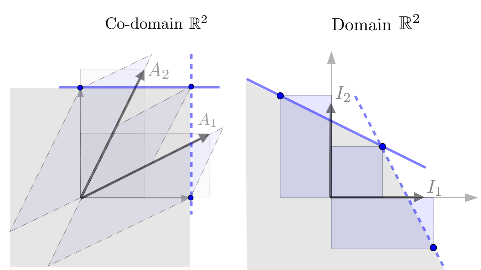
This paper is also meant to be part 1 of two part series; the second paper discusses matrix row geometry along a parallel track.[1]

[1]

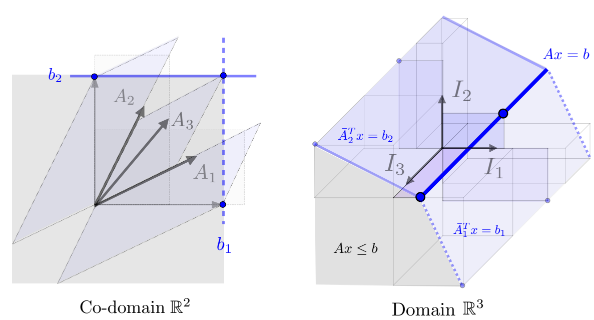
A. Inequality Column Geometry

The inequality version of the above constraints can be visualized in several ways. The most immediate and natural is to visualize the set on one side of the affine constraints given. Specifically for th 2×2 case given above, we give the image a





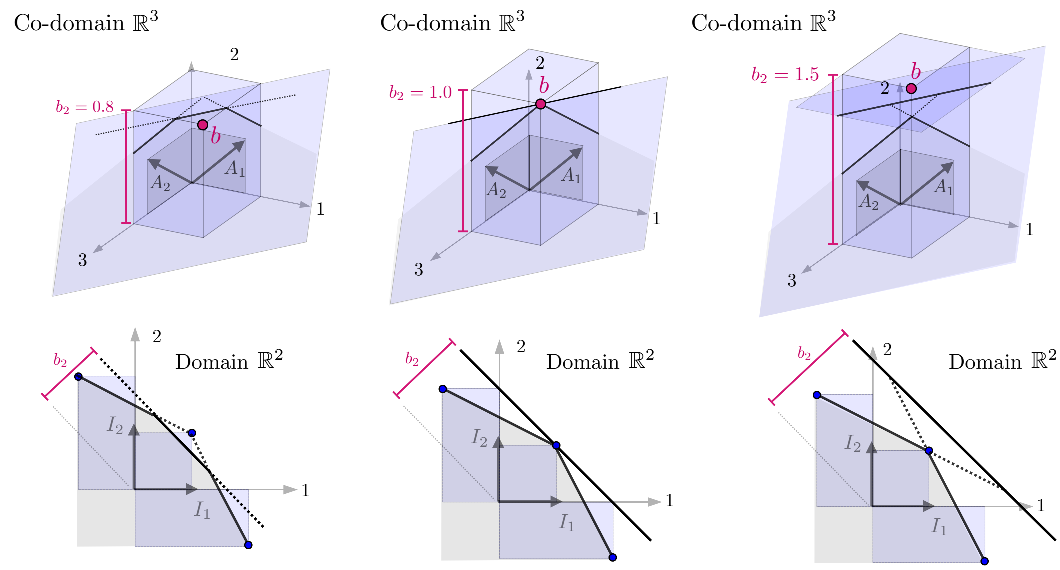
and preimage of inequality sets which are often more useful.



1) Slack Variable Representations: The general form for a polytope is given by

$$\{x \mid Ax = b,\ Cx \leq d\}$$

Often in practice, inequality constraints are dealt with analytically and in algorithms by adding *slack* variables



$s \in \mathbf{R}^m$. The original inequality constraint is then written as

$$\{x \mid Ax = b,\ Cx + s = d,\ s \geq 0\}$$

This has the benefit of "simplifying" the inequality part of the constraint to the form $s \geq 0$ at the cost of adding an extra affine constraint. This slack variable representation can be written in matrix form as

$$\left\{ \begin{bmatrix} x \\ s \end{bmatrix} \;\middle|\; \begin{bmatrix} A & 0 \\ C & I \end{bmatrix} \begin{bmatrix} x \\ s \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix},\ s \geq 0 \right\}$$

Geometrically, slack variables represent an arbitrary inequality of the form $Cx \leq d$ as the intersection of an affine constraint $Cx + s = d$ with the positive orthant (in s) $s \geq 0$. Assuming the matrix A is fat) the matrix

$$M = \begin{bmatrix} A & 0 \\ C & I \end{bmatrix}$$

will be fat and thus have a non-trivial nullspace regardless of the shape of C. If C is a fat matrix then, the inequality constraint can (likely) be satisfied with equality, ie. $Cx = d$, even with $s = 0$ and there may be a subspace of solutions as well (defined by the nullspace of C). If, however, C is tall, some non-zero slack s will be necessary to satisfy $Cx + s = d$.

A discussion of the nullspace of M is fruitful. We first consider the bottom rows $\begin{bmatrix} C & I \end{bmatrix}$. Taking the second set of columns (the identity block) to be a basis for the range space (which clearly it is), we get that elements in the nullspace of these rows can be written as

$$\begin{bmatrix} x \\ s \end{bmatrix} = \begin{bmatrix} I \\ -C \end{bmatrix} x$$

Here the vector on the right should indicate an arbitrary linear combination but the identity rows at the beginning indicate it will be x. We now consider hte first rows $Ax = 0$. Taking a basis for the nullspace of A given in the matrix N, we then have that $x = Nz$ and therefore the above equation becomes.

$$\begin{bmatrix} x \\ s \end{bmatrix} = \begin{bmatrix} I \\ -C \end{bmatrix} Nz = \begin{bmatrix} N \\ -CN \end{bmatrix} z$$

Note that this is the same characterization we would have gotten if we had first plugged in $x = Nz$, defined solutions to the equation $CNz + s = 0$

$$\begin{bmatrix} z \\ s \end{bmatrix} = \begin{bmatrix} I \\ -CN \end{bmatrix} z = \begin{bmatrix} I \\ -CN \end{bmatrix} z$$

and then mapped this set back into the (x, s) space instead of the (z, s) space

$$\begin{bmatrix} x \\ s \end{bmatrix} = \begin{bmatrix} N & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I \\ -CN \end{bmatrix} z = \begin{bmatrix} N \\ -CN \end{bmatrix} z$$

If we compute a basis for the nullspace of A and rewrite the affine constraint to have the form in it's nullspace form ($x = Nz + x_0$) then we can rewrite this as the image of a constraint on the variable z

$$\{z \mid CNz \leq \bar{d}\}$$

where $\bar{d} = d - Cx_0$ or in slack variable form as

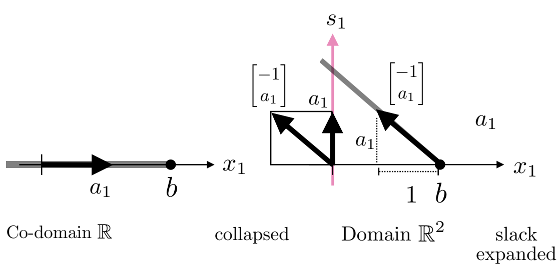
$$\{z \mid Ax = b,\ CNz + s = \bar{d},\ s \geq 0\}$$

We now give several low dimensional examples of this construction that are illustrative. Note the dimensions of the matrices in each each case.

We start with the most basic scalar case where $C \in \mathbb{R}^{1\times 1}$ for equations of the form

$$c_1 x_1 + s_1 = d_1, \quad s_1 \geq 0$$

s where $(x_1, s_1) \in \mathbb{R}^2$. Note that the inequality constraint defines a portion of $\mathbb{R}$ as illustrated in the figure below. The slack variable constraint relaxes the set to the (x_1, s_1) space $\mathbb{R}^2$. The original inequality is now the projection of this relaxed set onto the x_1 coordinate. This relationship is illustrated in the figure below.

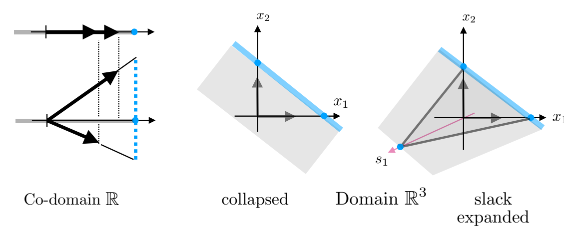


Note here that the dimension of the original inequality and the relaxed set are the same despite the fact that the second is in a larger ambient space. Note also the slack variable representatino of the set is the line $s = -c_1 x_1 + d_1$ in the (x_1, s_1) intersected with the positive half-space $s_1 \geq 0$. The slope of the line here is given $-c_1$. Note also that the affine space is orthogonal to the vector $\begin{bmatrix} 1 & -c_1 \end{bmatrix}$ which indeed forms a basis for the nullspace of the constraint matrix $\begin{bmatrix} c_1 & 1 \end{bmatrix}$. The value d_1 shifts the affine space horizontally along the x_1 axis.

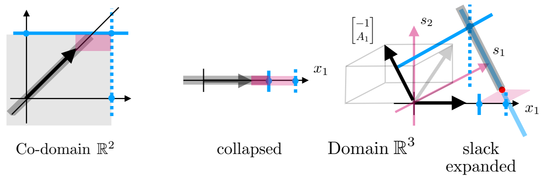
We now proceed to slightly richer examples. Specifically, we consider two of the most basic examples where C is not square. We start with the case where C is fat; $C = \begin{bmatrix} c_1 & c_2 \end{bmatrix} \in \mathbb{R}^{1\times 2}$ and $d_1 \in \mathbb{R}$. In this case the constraints can be represented as

$$\begin{bmatrix} c_1 & c_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + s_1 = d_1$$

Here $x \in \mathbb{R}^2$ is a 2D vector with one constraint. Technically, the "column" geometry of C should really be expressed on the individual number line. However, as per the discussion above, we can expand the points on the number line with an extra dimension in order to see the pre-image better. The set $Cx = d$ is a 1D affine space shown in the image below. Note that since C is fat there is always a solution to the equality constraint; indeed there is one degree of freedom provided by the nullspace of C. Note the intercepts at $x_1 = d_1/c_1$ and $d_1/c_2 = x_2$ and that the affine space is orthogonal to the vector $\begin{bmatrix} c_1 & c_2 \end{bmatrix}$ as expected. Adding a slack variable increases the ambient dimension to 3D and the set is now expressed as the affine space given above intersected with the half space $s_1 \geq 0$. Visually, we can think about the half space tilting up from the x-plane through the intercept $s_1 = d_1$.



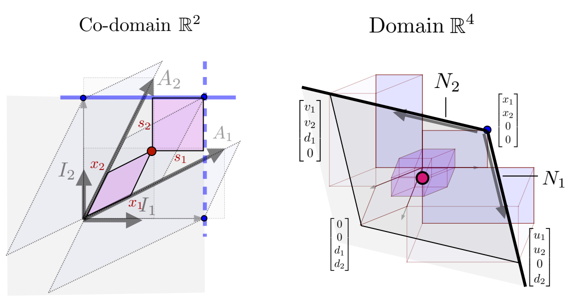
We now proceed to the case where C is tall; $C \in \mathbb{R}^{2\times 1}$. In this case $x = x_1$ is simply a 1D variable and the constraints provide two inequality constraints on the same variable. Often only one of them will be relevant. For example, take $c_1 = c_2 = 1$ and $d_1 = 1$ and $d_2 = 2$. In this case the constraints become $x_1 < 1$ and $x_1 < 2$. Only the first constraint here is relevant. Algebraically, this corresponds to the fact that since C is tall, there may not be a solution to the equality $Cx = d$. Here the slack variables are critical for the constraint $Cx + s = d$ to be satisfied with equality as the slack variables make up any "slack" between a row of Cx and that element of d. The geometry of this slack is illustrated in the figure below. The inequality representation in the domain is then represented on the right. When we add in the slack variables, the space expands to $\mathbb{R}^3$. Unlike before we added one constraint (and thus one slack variable dimension) to a two dimension space, here we add two slack variable dimensions to a one dimensional space. The affine space has a 1D nullspace spanned by $\begin{bmatrix} -1 & c_1 & c_2 \end{bmatrix}^T$. The feasible set is the intersection of this 1D subspace with the orthant $s = (s_1, s_2) \geq 0$. Note here some of the intercepts. If $x_1 = 0$, then $s = d$. There is no solution where $s = 0$; however when one of the inequality constraints is met we can set that element of s equal to 0.



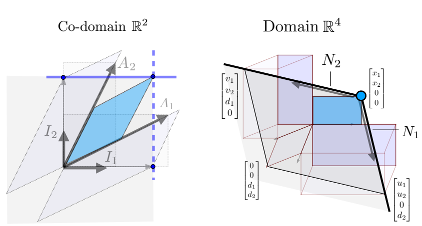
We lastly turn to the case where $C \in \mathbb{R}^{2\times 2}$

$$C = \begin{bmatrix} | & | \\ C_1 & C_2 \\ | & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

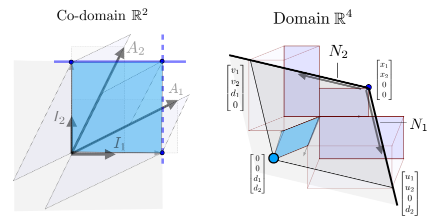
Here the original variables are 2D $x \in \mathbb{R}^2$ and two constraints correspond to two slack variables $s \in \mathbb{R}^2$. The slack variables lift the original inequality set to the intersection between the orthant $s \geq 0$ and the affine space defined by the matrix $\begin{bmatrix} C & I \end{bmatrix} \in \mathbb{R}^{2\times 4}$. For a given linear combination of the columns of C (defined by x) the slack variables make up the difference to the point d as shown in this figure.



Since C is invertible, if $x = C^{-1}d$, then $s = 0$, ie. no slack variables are needed since both inequalities are satisfied with equality.



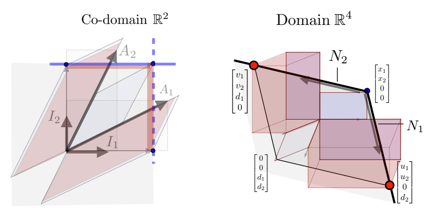
If $x = 0$, then the constraints must be satisfied only with slack variables and $s = d$.



Finally, if we let

$$x' = C^{-1}\begin{bmatrix} 0 \\ d_2 \end{bmatrix}, \qquad x'' = C^{-1}\begin{bmatrix} d_1 \\ 0 \end{bmatrix},$$

then Cx' satisfies the first inequality and the slack variable must make up the second constraint, ie. $s_2 = d_2$ and Cx'' satisfies the second inequality and the slack variable must make up the slack in the first constraint ie. second constraint, ie. $s_1'' = d_1$.



Each of these points (in the domain)

$$\begin{bmatrix} x_1 \\ x_2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ d_1 \\ d_2 \end{bmatrix}, \begin{bmatrix} x_1' \\ x_2' \\ 0 \\ d_2 \end{bmatrix}, \begin{bmatrix} x_1'' \\ x_2'' \\ d_1 \\ 0 \end{bmatrix}$$

and their images in the co-domain are illustrated in Figure XXX. Note this space is 4D and thus difficult to visualize. The vertical plane is the first two coordinates (the x-coordinates). The slack variable axes are orthogonal to both the x coordinates and each other tho this is difficult to visualize. The affine space is a 2D subspace and the points given above are actually the corners of a parallelogram in this 2D space. The sides of the parallelogram are defined by the differences of the corners and provide a basis for the nullspace which can be written in the columns of a 4×2 matrix

$$\begin{bmatrix} x_1' & x_1'' \\ x_2' & x_2'' \\ 0 - d_1 & d_1 - d_1 \\ d_2 - d_2 & 0 - d_2 \end{bmatrix} = \begin{bmatrix} x_1' & x_1'' \\ x_2' & x_2'' \\ -d_1 & 0 \\ 0 & -d_2 \end{bmatrix}$$

Note here that by definition we have that

$$C\begin{bmatrix} x_1' & x_1'' \\ x_2' & x_2'' \end{bmatrix} = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$$

From this constructions we can write $\begin{bmatrix} x' & x'' \end{bmatrix} = C^{-1}\mathbf{dg}(d)$ and

$$\begin{bmatrix} x_1' & x_1'' \\ x_2' & x_2'' \\ -d_1 & 0 \\ 0 & -d_2 \end{bmatrix} = \begin{bmatrix} C^{-1} \\ -I \end{bmatrix} \mathbf{dg}(d) = \underbrace{\begin{bmatrix} I \\ -C \end{bmatrix}}_{N} C^{-1}\mathbf{dg}(d)$$

whose columns are simply (linearly independent) linear combinations of the columnns of the matrix N which is the natural basis we would construct from our discussion of nullspaces above. We can also show that this set is indeed a parallelogram (and thus all four points lie in the same 2D space by adding the two sides to one corner to get the across corner. Explicitly, we know that $Cx = d = Cx' + Cx''$ and it follows that $x = x' + x''$. Therefore we have that

$$\begin{bmatrix} 0 \\ 0 \\ d_1 \\ d_2 \end{bmatrix} + \begin{bmatrix} x_1' \\ x_2' \\ -d_1 \\ 0 \end{bmatrix} + \begin{bmatrix} x_1'' \\ x_2'' \\ 0 \\ -d_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \\ 0 \end{bmatrix}$$

Remark 1. *These slack variable examples are subtle and*

II. Row Geometry

Defining columns of a matrix inevitably defines rows and thus offers another natural geometry perspective on matrices. Here we represent a matrix in the form

$$A = \begin{bmatrix} - & \bar{A}_1^T & - \\ & \vdots & \\ - & \bar{A}_m^T & - \end{bmatrix}$$

Multiplying this matrix by a vector Ax takes the inner product of the vector with each row

$$Ax = \begin{bmatrix} - & \bar{A}_1^T & - \\ & \vdots & \\ - & \bar{A}_m^T & - \end{bmatrix} \begin{bmatrix} | \\ x \\ | \end{bmatrix} = \begin{bmatrix} \bar{A}_1^T x \\ \vdots \\ \bar{A}_m^T x \end{bmatrix}$$

Inner products naturally encode the idea of projection and thus loosely speaking we are can determine the coordinate of the output vector Ax by "projecting" x onto each row. This process can be visualized using the various innner product visualizations presented earlier. We already discussed the linear combination inner product visualization method for this in our discussion of column geometry (since it is naturally leveraging column geometry to see inner products of rows). Here we will focus the bulk of our efforts on visualizing these inner products when the rows are presented as vectors in the domain, ie. using the "projection method".

A. Set Transformations

1) Image of Domain Sets:

2) Pre-image of Co-Domain Sets:

B. Subspace Geometry

Subspaces can be represented algebraically as the nullspace of a matrix or as the range of matrix. For a matrix $A \in \mathbb{R}^{m\times n}$ with rank k, we have

Nullspace representation	$\{x \in \mathbb{R}^n \mid Ax = 0\}$
Range space representation	$\{x \in \mathbb{R}^n \mid x = Nz,\ z \in \mathbb{R}^k\}$

where $N \in \mathbb{R}^{n\times k}$ forms a basis for the nullspace of A. There are many methods for computing the basis N such as Gaussian elimination or singular value decomposition (see XXX for details). We sketch the details for one method as it will be critical in understanding the nullspace relative to the column geometry of A. If we select k linearly independent columns of A (wlog assume they are the first k columns) we can write $A = \begin{bmatrix} A' & A'' \end{bmatrix}$ with $A' \in \mathbb{R}^{m\times k}$ and A'' containing the remaining $n-k$ columns. Since A has rank k, we can write each column of A'' as linear combinations of the columns of A', ie. $A'' = A'B$ for some matrix $B \in \mathbb{R}^{k\times n-k}$. We then have that $A = A'\begin{bmatrix} I & B \end{bmatrix}$. We can see immediately then that if

$$N = \begin{bmatrix} B \\ -I \end{bmatrix}$$

then $AN = 0$. With a little more work, we can show that the columns of N form a basis for the nullspace. The identity block proves the linear independence of the columns. We can also show explicitly that any element in the nullspace of A is in the span of N. This fact relies on the linear independence of the columns of A'. Explicitly these two proofs are given by

$$\text{LIN IND:} \qquad Nx = 0 \Rightarrow \begin{bmatrix} Nx \\ -x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow x = 0$$

$$\text{SPAN:} \qquad Ax = 0 \Rightarrow A'\begin{bmatrix} I & B \end{bmatrix} \begin{bmatrix} x' \\ x'' \end{bmatrix} = 0$$

$$\Rightarrow x' = -Bx'' \Rightarrow x = \begin{bmatrix} B \\ -I \end{bmatrix} (\text{-}x'')$$

where the beginning the second line depends on the columns of A' being linearly independent. The geometry

A subspace

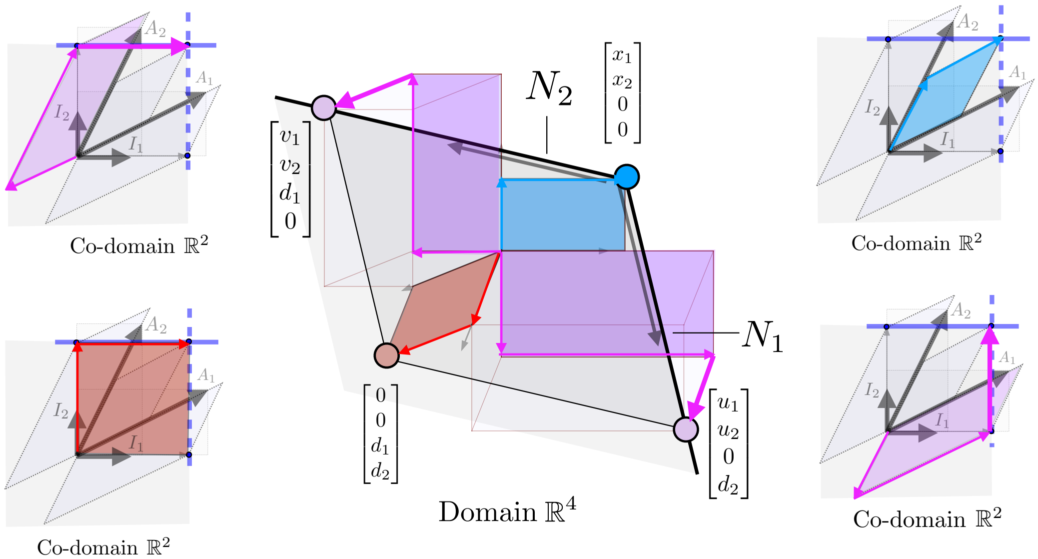
Given these

The geometry of this construction is

There are two natural representations of subspaces: the affine representation and

1) Row Geometry: A subspace ahs tw

Traditional



A_2
A_1
I_2
I_1
Co-domain $\mathbb{R}^2$
A_2
A_1
I_2
I_1
Co-domain $\mathbb{R}^2$
N_2
$\begin{bmatrix} x_1 \\ x_2 \\ 0 \\ 0 \end{bmatrix}$
$\begin{bmatrix} v_1 \\ v_2 \\ d_1 \\ 0 \end{bmatrix}$
N_1
$\begin{bmatrix} 0 \\ 0 \\ d_1 \\ d_2 \end{bmatrix}$
$\begin{bmatrix} u_1 \\ u_2 \\ 0 \\ d_2 \end{bmatrix}$
Domain $\mathbb{R}^4$
A_2
A_1
I_2
I_1
Co-domain $\mathbb{R}^2$
A_2
A_1
I_2
I_1
Co-domain $\mathbb{R}^2$