Column Geometry Visualizations

Dan Calderone

***Abstract*—This tutorial paper gives basic visualizations for matrix column and row geometry. Basic vector visualization as well inner product visualization techniques are discussed first. The body of the paper is divided into a discussion of matrix column geometry and matrix row geometry. Each section discusses visualization of images and pre-images of sets in the domain and co-domain relative to each geometry. Various domain and co-domain subspaces are discussed with specific focus given to the range and nullspaces of the matrix and corresponding transpose. Affine spaces are discussed and then linear inequalities and polytope geometry. Particular focus is 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 and rows 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.

The initial introductory section focuses on basic vector visualization techniques. We present two different techniques, one we refer to as the *orthogonal (or spatial) axis* representation (the traditional representation) and what we call the *parallel axis* representation. We discuss these representations in two, three, and higher dimensions as well as limits of visualization. We also several methods for visualizing linear combinations and inner products which will be foundational for the rest of the paper and then continue on to matrices.

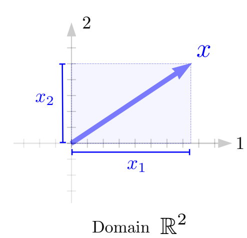
In many ways, the column geometry of a matrix is the natural starting point for our discussion. The columns of a matrix define where the standard basis vectors (and thus the axes) in the domain map to in the co-domain. For this reason, visualizing the columns of a matrix provides a natural way to visualize the action of the transformation. Defining columns also defines rows which tell how a particular axis in the output space is affected by vectors in the domain; in particular this is given by taking an inner product of a particular vector with the row.[1]

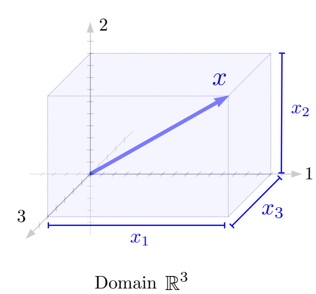
II. Vector Visualizations

Finite dimensional vectors are represented as a string of digits, each of which gives a displacement relative to an axis. Visual representations of vectors show these displacements in various ways. We will focus on two methods which we will refer to as the *orthogonal (or spatial) axis* representation and the *parallel axis* representation.

A. Spatial (Orthogonal) Axes

The natural (spatial) way to represent the various displacements is along axes that are orthogonal to each other visualized in 2D and 3D here.

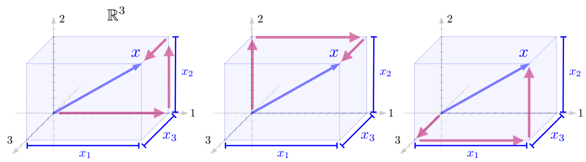




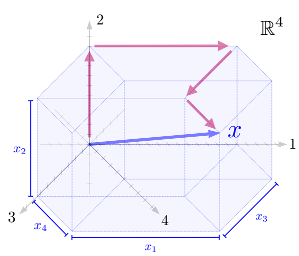
Cubes (or rectangles) can be used to visualize each individual coordinate of the vector. The length of each cube edge shows the size of each coordinate. Note also that we can visualize building up a vector one coordinate at a time as walking along edges of the cube from the origin to the tip. Note there are multiple different paths

[1]©Dan Calderone, September 2022

that all result in reaching the tip corresponding to the fact that coordinates can be added in any order.

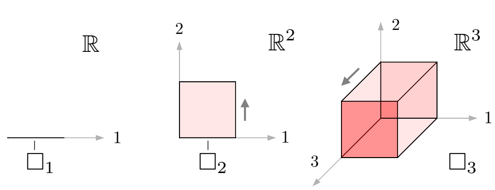


Our brains are highly adapted for visualizing vectors orthogonally in two and three dimensions. If we wish to use the orthogonal representation for vectors in dimensions higher than three, the best we can do is to draw a 2D projection of the higher dimensional vectors. One projection method can be achieved by simply drawing a direction (in 2D) for each axis and then simply showing displacements along these axes. This process is illustrated here for a vector in $\mathbb{R}^4$.

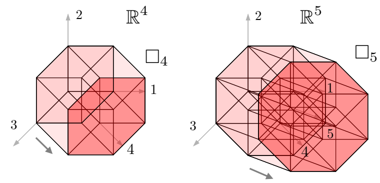


Since there is only room for two orthogonal axes in 2D, we must imagine that the axes we draw are actually orthogonal.

Remark 1. *The following exercise is useful for visualizing hypercubes in higher dimension. A 1D cube is a basic line segment along the interval* $[0, 1]$*. We can obtain a 2D cube, a square, by sweeping that interval along a 2nd axis. Sweeping the square along a 3rd axis produces a 3D cube.*



Higher dimensional cubes can be produced by continuing this process. A 3D cube swept along a 4th axis gives a 4D-hypercube; a 4D-hypercube swept along a 5th axis produces a 5D-hypercube; etc.

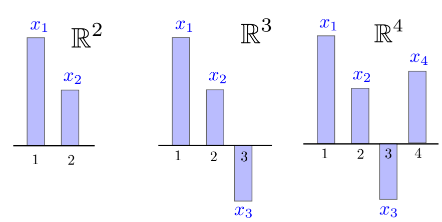


Remark 2. *When projecting higher dimensional shapes onto 2D images, a certain amount of information, the "depth" direction(s), in the image get lost. If we're drawing a 3D vector, depth is one dimensional (out of the page). If we're drawing a 4D vector, depth is 2-dimensional; for a 5D vector, depth is 3-dimensional, etc. Any intuition derived from projections of higher dimensional sets should be verified with rigorous proof.*

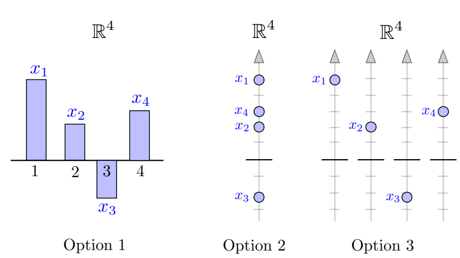
Referring to these visualization techniques as an "orthogonal" representation is somewhat of a misnomer because the axes are only truly orthogonal in 2D. We use this term to reference the fact that we are conceptualizing the axes as orthogonal in some higher dimensional space even though we can only view 2D projections. The term *spatial* representation is perhaps more accurate.

B. Parallel Axes

Another less traditional way to represent vectors is to place the axes parallel to each other and show each coordinate displacement in the same direction. We will often use rectangles of the appropriate heights to visualize this type of representation as demonstrated here for vectors in 2D, 3D, and 4D.



Note that for negative coordinates the rectangles extend down from the zero level. Other options are possible particularly visualizing coordinates as points on parallel axes or on the same axis.

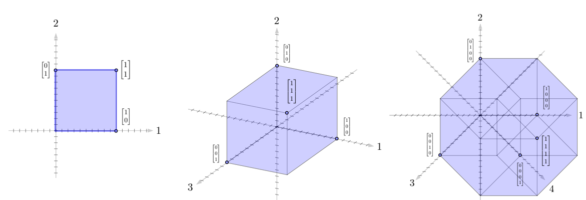


This parallel axis representation does not suffer from the pathology of depth; however this is simply because it does not seek to leverage our 3D spatial intuition. For example, even in 2D or 3D it is quite difficult to see immediately that two vectors are orthogonal in their parallel axes representation. (Readers are encouraged to try this.) This parallel axis representation will arise in several ways in our treatment of these subjects. First, when we represent columns of a matrix using a spatial representation, the rows naturally appear in a parallel representation along each axis (and vice versa if the rows are represented spatially, the columns appear in a parallel representation.) While not used significantly in this paper, this is an interesting fact that the authors hope to explore in future work. Second, a hybrid parallel-spatial geometric technique can be quite useful in visualizing inner products (discussed below).

III. Basic Geometric Sets

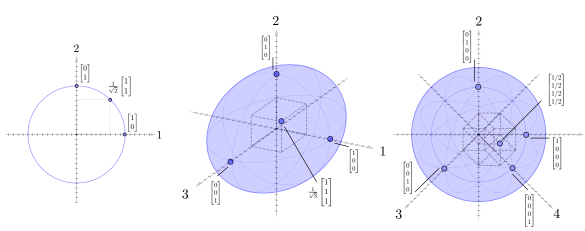
We briefly discuss several geometric sets that we will refer to later. For each set, we label several examples of interesting points. We have already discussed unit cubes briefly. Similarly, let $\square_n$ represent the n-dimensional unit cube.

$$\square_n = \left\{ x \in \mathbb{R}^n \;\middle|\; \mathbf{0} \leq x \leq \mathbf{1} \right\}$$



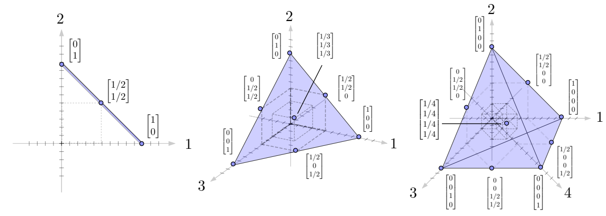
Let $\bigcirc_n$ represent the n-dimensional unit sphere.

$$\bigcirc_n = \left\{ x \in \mathbb{R}^n \;\middle|\; \|x\|_2 = 1 \right\}$$



Note that depending on the axes are drawn a unit circle may not appear perfectly circular. Let Δ_n represent the n-dimensional simplex.

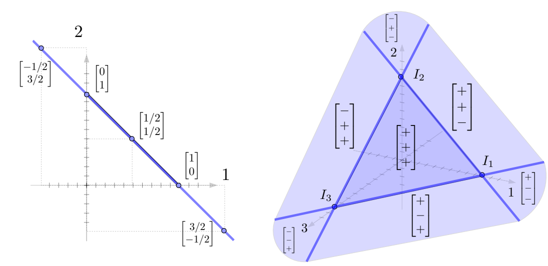
$$\Delta_n = \left\{ x \in \mathbb{R}^n \;\middle|\; \mathbf{1}^T x = 1,\; x \geq 0 \right\}$$



Simplexes are often used to represent probability distributions on finite sets. We note that if we remove the constraints that each element $x_i \geq 0$, we get a a particular $n-1$-dimensional affine space

$$\ell_n = \left\{ x \in \mathbb{R}^n \;\middle|\; \mathbf{1}^T x = 1 \right\}$$

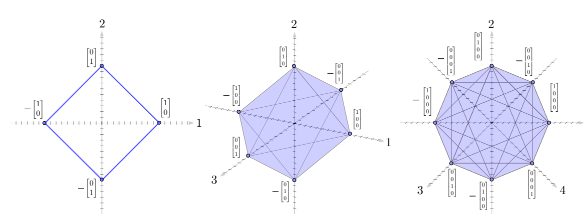
For $n = 2$ this is a line through the points I_1, I_2; for $n = 3$ this is a plane through the points I_1, I_2, I_3. Later on, this set will be useful in defining many affine spaces via transformation through a matrix.



It will also be useful to represent unit balls of other norms. The 2-norm unit ball is simply the Euclidean unit sphere above. Let $\Diamond_n$ represent the n-dimensional unit 1-norm ball.

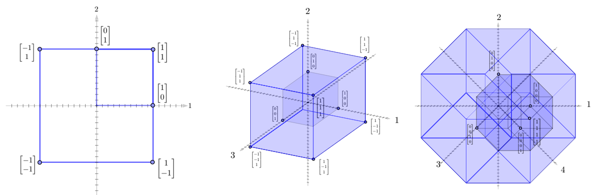
$$\Diamond_n = \left\{ x \in \mathbb{R}^n \;\middle|\; \|x\|_1 = 1 \right\}$$

Note that the "faces" of the 1-norm ball are simplicies with different sign patterns.



Finally, let $\square_n^\infty$ represent the n-dimensional unit ∞-norm ball.

$$\square_n^\infty = \left\{ x \in \mathbb{R}^n \;\middle|\; -\mathbf{1} \leq x \leq \mathbf{1} \right\}$$



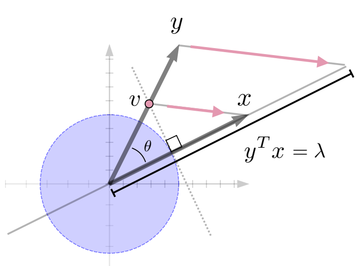
Note that the unit ∞-ball is substantially larger than the unit cube though it has the same shape. The unit cube always has unit volumew regardless of the dimension. The volume of the unit ∞-ball grows rapidly as 2^n.

IV. Inner Products

Visualizing inner products, $y^T x$ for $x, y \in \mathbb{R}^n$ is critical to any geometric visualization. We detail two methods here: one based on the spatial axis representations of vectors and one based on a hybrid parallel-spatial axis representation. Note that while we will present these techniques with the two vectors y and x playing different roles, inner products are symmetric and the roles of these vectors can always be reversed.

A. Spatial Visualization

The first technique is a somewhat traditional approach to inner product visualization where an inner product is thought of as a projection from one vector onto another. Consider two vectors $x, y \in \mathbb{R}^n$. First define the unit vector in the x direction and an $n-1$ dimensional affine space tangent to the unit sphere at this point. Next, define a vector v in the y-direction on this plane. If we drag this point v to the tip of x (and move y along with it), then y moves to a vector with length $y^T x$. This process is illustrated (for $y, x \in \mathbb{R}^2$) here.

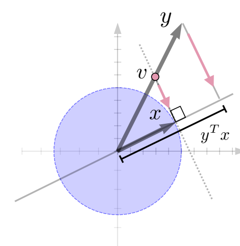


A brief algebraic justification for this is warranted.

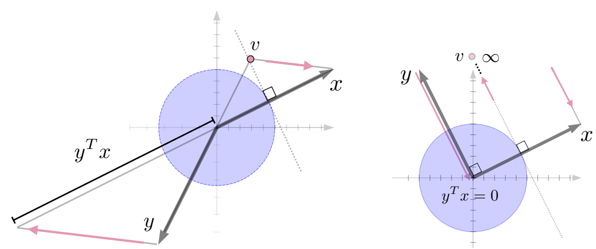
Proof: From properties of similar triangles and the definition of v we can obtain that

$$\frac{\lambda}{\|x\|_2} = \frac{\|y\|_2}{\|v\|_2}, \qquad \|v\|_2 \cos\theta = 1$$

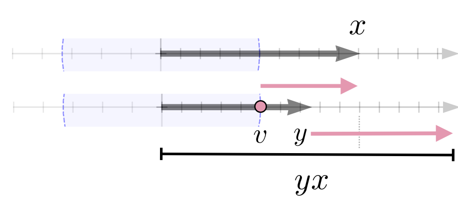
It follows that $\lambda = \|y\|_2 \|x\|_2 \cos\theta = y^T x$ If x is a unit vector, ie. $\|x\|_2 = 1$, this visualization becomes the traditional visualization of an inner product as a projection. One can think of this technique as extending the projection idea to the case when x and y are both non-unit vectors.



The reader is encouraged to experiment with other values of x and y. Here we show the case where the inner product is negative and the important limiting case where y and x are orthogonal ($y^T x = 0$) and the point v goes to ∞, ie. y does not intersect the tangent plane.



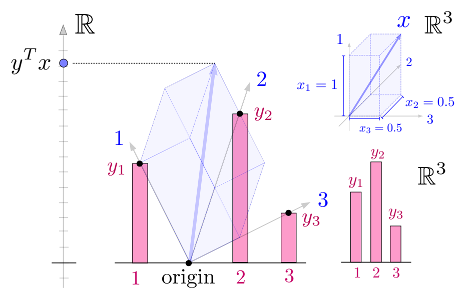
We also note that this technique reduces to perhaps the most natural way to visualize scalar multiplication. For two values $x, y \in \mathbb{R}$ on a number line, one can think of taking the product yx as stretching a unit value for y to be the number x, ie. treating x as the “units” for y. This “rescaling” of the y number line (by x) moves y to yx.



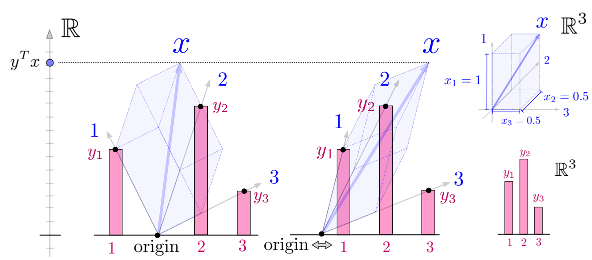
B. Parallel-Spatial

The second inner product visualization technique combines the parallel axis and spatial axis representations of vectors. This technique is subtle and actually quite powerful and is closely related to the idea of column geometry visualization which is at the heart of this paper.

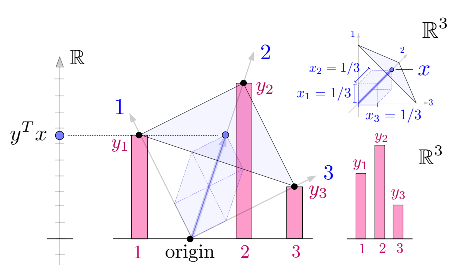
For this technique, we will one vector, y, in it's parallel representation, and the other vector, x, in it's spatial representation. We illustrate this for $y, x \in \mathbb{R}^3$. First, draw y as a set of heights (or parallel displacements). Next, draw axes with the tips of unit vectors at the ends of each coordinate values of y and place the origin anywhere along the zero-value line. Finally, draw the spatial representation of x relative to this axis. The height of the corresponding point (or, more generally, the displacement along the parallel axis direction) is $y^T x$. This technique is illustrated in the figure below.



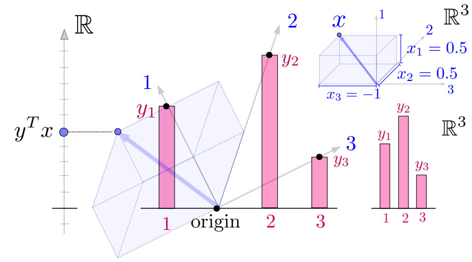
It should be noted that the location of the origin and actually all the horizontal values in this picture are irrelevant to the final result and can be shifted freely. We will often take advantage of this to improve the legibility of various illustrations. Again, the reader is encouraged to try this.



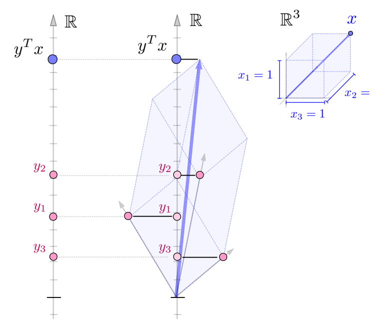
This technique is particularly useful when the vector x is a convex combination, ie. $\mathbf{1}^T x = 1,\ x \geq 0$. In this case, the image of x in the hybrid visualization will appear in the convex hull of of the heights of the $y_i's$ as illustrated here. This image and then the value $y^T x$ are easy to see.



On the other hand, the visualization can be more difficult to use when the values of x are large or negative. It still works, however.



It is worth noting that the worst option for the horizontal position of the elements of y is to line them all up on the same axis. Providing variation in the horizontal configuration allows us to differentiate the vector elements and "see" the spatial representation of x better. In fact, many times even if the points naturally lie on the same axis, we will simply "bump" each of them off the axis an arbitrary (but different) amount. When we project the result back onto the original axis to visualize $y^T x$ we are removing the horizontal information that we added initially.

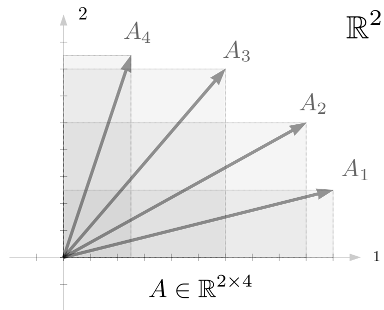


Also it should be noted that there is nothings special about "horizontal" and "vertical" in these examples and there roles can be swapped with the same intuition. However, it is important that the "perturbation" direction is orthogonal to the primary direction.

V. Column Geometry

The columns of a matrix $A \in \mathbb{R}^{m \times n}$ are vectors in the co-domain of the linear map

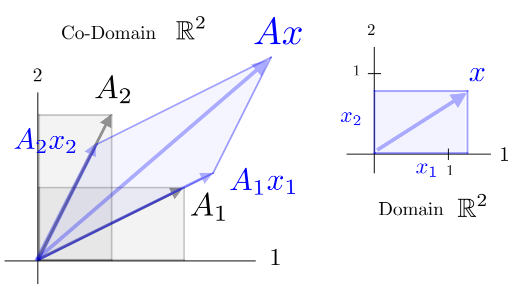
$$A = \begin{bmatrix} | & & | \\ A_1 & \cdots & A_n \\ | & & | \end{bmatrix}$$



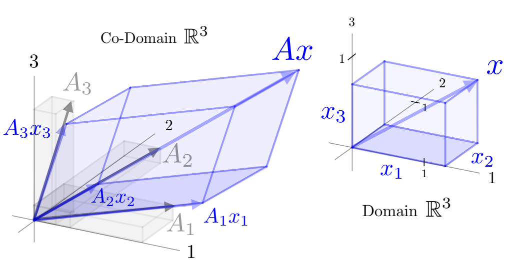
Each individual column $A_j \in \mathbb{R}^m$ tells where the jth standard basis vector (in the domain) gets mapped under the transformation. Explicitly $AI_j = A_j$. We can see where a vector $x \in \mathbb{R}^n$ in the domain gets mapped by breaking up x into a linear combination of standard basis vectors (ie. $x = I_1x_1 + \cdots + I_nx_n$), transforming each standard basis vector to the appropriate column, and then recombining. Algebraically, this is given by

$$\begin{aligned} Ax &= A\big(I_1x_1 + \cdots + I_nx_n\big) \\ &= AI_1x_1 + \cdots + AI_n \\ &= A_1x_1 + \cdots + A_nx_n \end{aligned}$$

Graphically, we illustrate this process below for matrices $A \in \mathbb{R}^{2\times 2}$

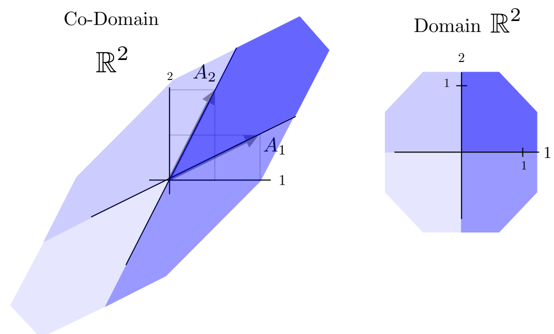


and also $A \in \mathbb{R}^{3\times 3}$.



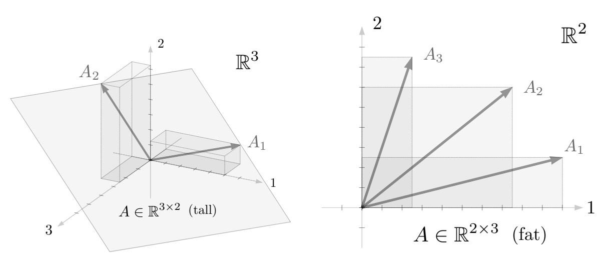
As in our discussion of vectors, the various routes from the origin to the tip of the vector along edges of the hypercube relate to the different order the scaled columns can be added in. Note also that if we take $A = I$, we simply get the vector itself back.

We can therefore "see" vectors in the domain by squinting our eyes and visualizing the axes of the domain ($\mathbb{R}^n$) positioned relative to the columns of A as illustrated here in the 2×2 case.

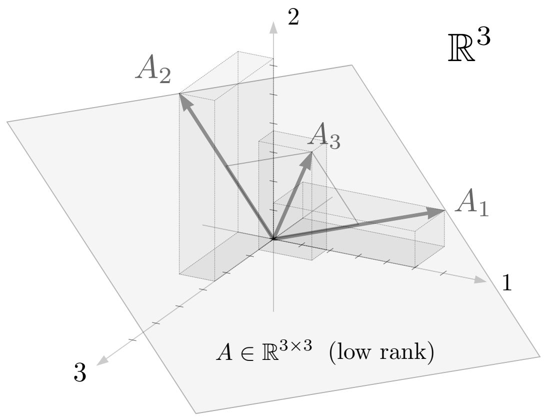


We will use this technique ad nauseam in what follows so it is worth getting comfortable with it.

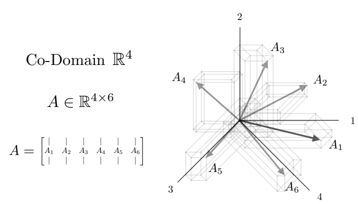
Note here there is nothing that immediately requires that A be square. For tall matrices (fewer columns than dimensions) the reachable points will be a subset of the co-domain; for fat matrices (more columns than dimensions) there will be redundant ways to reach any point in the co-domain. We give two examples here, a tall matrix $A \in \mathbb{R}^{3\times 2}$ and a fat matrix $A \in \mathbb{R}^{2\times 3}$



It is also possible to have columns that only span a proper subspace of $\mathbb{R}^m$ but are not linearly independent from each other. This is true for any matrix that is both column and row rank deficient. We illustrate one possibility here for $A \in \mathbb{R}^{3\times 3}$ with rank 2. Here column 3, A_3 is linearly dependent on A_1 and A_2.



For more than two or three columns, the space we will be visualizing is high dimensional as in the vector illustrations in Figure XXX (with all the associated problems of depth). The co-domain of the map is ambient vector space that the columns live in and is thus easy to visualize. Depending on how ambitious we are, this space may be high dimensional as well. We illustrate the column geometry of a 4×6 matrix in the figure below in order to give a flavor for what such an attempt might look like. Again, "depth" would prove a major problem in this case both in the domain and co-domain and so such pictures are limited in their usefulness.

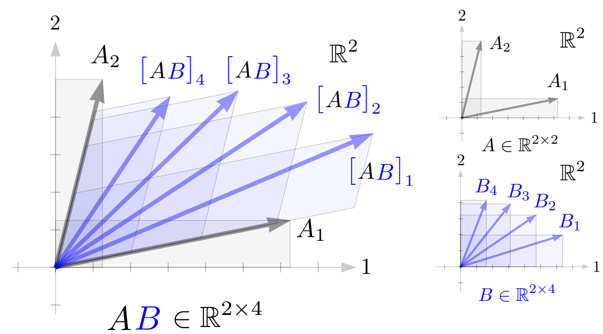


Remark 3. *If we want to focus in on a particular coordinate of the output, we project the output vector onto that particular axis. Algebraically, here we are considering the inner product between x and that particular row of A, $\bar{A}_j^T x$ We note that here if we just focus at the distance of the output along a particular axis we are actually using the linear combination visualization for inner products presented before to visualize $\bar{A}_i^T x$. Note that as discussed in the inner product visualization section the position of the columns in the other coordinates $i' \neq i$ will not actually affect the length of the output along the ith direction. We can think of visualize Ax as actually using the linear combination innner product visualization method on all axes at once as illustrated in Figure XXX. This is perhaps a backwards view because this linear combination inner product visualization technique is actually leveraging the power of column geometry to visualize inner products (instead of vice versa) but the connection is interesting and worth noting. We will discuss row geometry from alternative perspectives in the second half of the paper.*

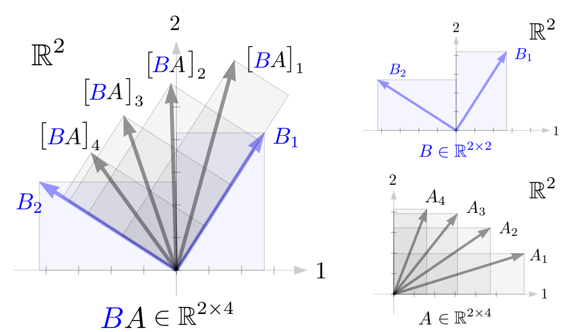
A. Matrix Multiplication for Column Geometry

1) Right Multiplication : When we right multiply a matrix $A \in \mathbb{R}^{m\times n}$ by a matrix $B \in \mathbb{R}^{n\times p}$, each columns of B take linear combinations of the columns of A. Algebraically, this corresponds to treating each column of B as a separate vector that gets multiplied by A, ie.

$$AB = A\Big[B_1 \;\cdots\; B_p\Big] = \Big[AB_1 \;\cdots\; AB_p\Big]$$



2) Left Multiplication: Left multiplication of A by B (of appropriate dimensions) transforms each individual column by the same transformation (B). separately as illustrated here when B is a simple rotation matrix.



We use a simple visual example for B (a rotation) to clearly illustrate that the transformation is applied to each column of A, but of course more complicated transformations of B would work.

Remark 4. *For a complicated matrix B, it's probably easiest to understand how it transforms A by thinking of how columns of A are represented relative to B, ie.*

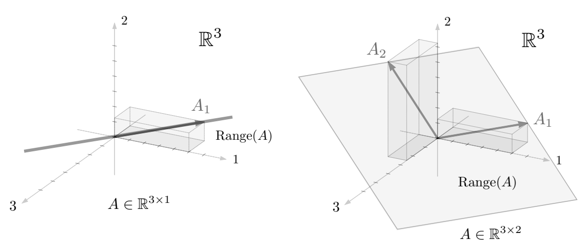
thinking about B being right multiplied by A. Of course this is fine and whether the product BA is seen as A left multiplied by B or B right multiplied by A is simply interpretation.

We now turn our focus to visualizing sets in the co-domain and the domain relative to the column geometry of a matrix. We start with the co-domain since column geometry provides a a natural way to visualize images of (domain) sets transformed by the matrix. The bulk of this section will focus on visualizing domain set, ie. pre-images of (co-domain) sets. These visualizations will be significantly more subtle and in some ways even more fruitful.

B. *Co-Domain Sets*

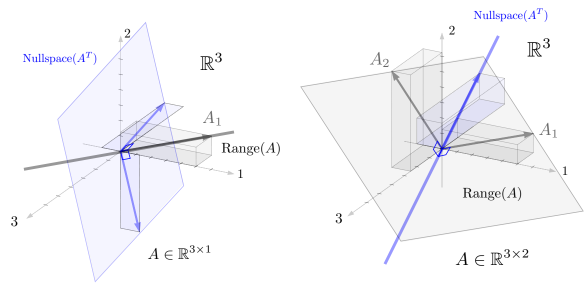
Column geometry is quite natural for visualizing sets in the co-domain. In particular, we will focus on images of sets (in the domain) transformed by the matrix.

1) Range: The image of the entire domain through the matrix is called the range. Visualizing the range is quite natural in terms of column geometry since it is simply the span of the columns of the matrix. Note for a matrix $A \in \mathbb{R}^{m \times n}$ with rank k the dimension of the range is k, the span of k linearly independent columns. If $k \geq m$, then the span is the entire co-domain. We illustrate the range of two matrices with co-domain of $\mathbb{R}^3$, $A \in \mathbb{R}^{3\times1}$ and $\mathbb{R}^{3\times2}$.



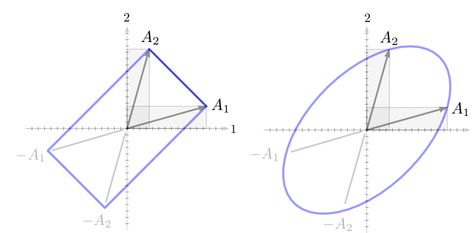
The discussion above gives several other good examples of matrix ranges.

2) Nullspace of A^T: The orthogonal complement to the range of $A \in \mathbb{R}^{m\times n}$ is the set of vectors with no component in the range. This subspace corresponds with the nullspace of A^T, ie. is orthogonal to all the columns of A. We illustrate this here for the range examples above. Note that the dimension of the nullspace of A^T is $m - k$, the difference between m, the dimension of the co-domain m and the rank k. As a result, if $m = k$ then the nullspace has dimension 0 and is simply the 0 vector.

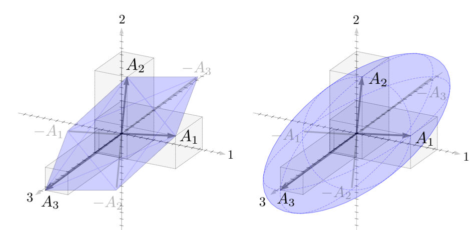


We now proceed to images of basic sets. Note that all the images that follow will be subsets of the range.

3) Basic Set Images: The images of several basic sets, the 1-norm ball and the 2-norm ball are given below for $A \in \mathbb{R}^{2\times2}$

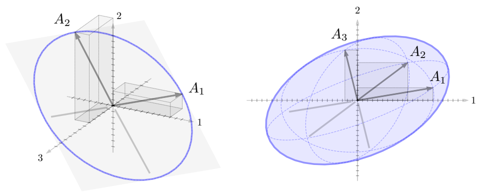


and $A \in \mathbb{R}^{3\times3}$

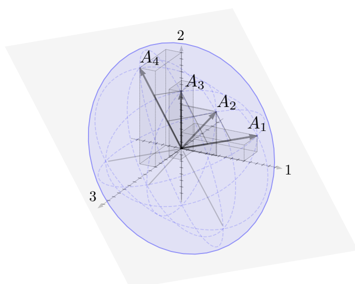


Note that the simplex is highlighted as well. These sets were chosen as they show the geomery of each orthant (in the domain) under the transformation. The location of each orthant should be noted by the reader.

We now show several examples of the image of the 2-norm ball for rank-deficient matrices. The first two matrices are non-square but full column or row rank depending on context. The first matrix $A \in \mathbb{R}^{3\times2}$ is tall and thus the columns do not span the full co-domain and as a result the output image is flat relative to the co-domain. The second matrix $A \in \mathbb{R}^{2\times3}$ is fat. Some of the directions in the domain get mapped to the same point in the domain, ie. some information is lost (literally flattened) in the transformation.



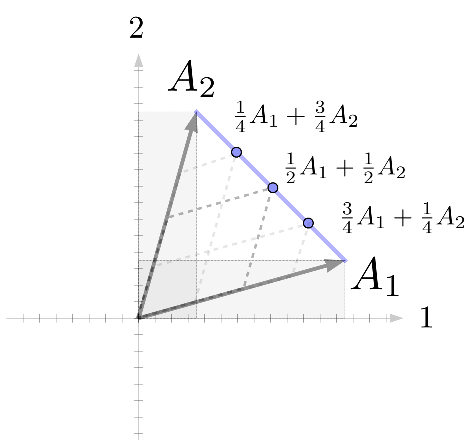
Finally, we show a matrix that is both column and row rank deficient, $A \in \mathbb{R}^{3\times 4}$ with rank 2. Here the matrix flattens two dimensions of the 4D set and then presents them in a 2D subspace of the 3D domain.



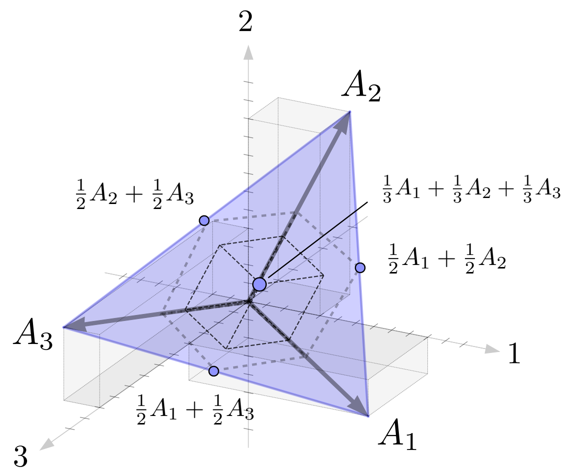
4) Convex Hulls: We now focus specifically on the image of the simplex, ie. the convex hull of the columns of A. We will denote this set $A\Delta$ as is fitting

$$
\begin{aligned}
A\Delta &= \left\{ Ax \in \mathbb{R}^m \;\middle|\; x \in \Delta_n \right\} \\
&= \left\{ Ax \in \mathbb{R}^m \;\middle|\; \mathbf{1}^T x = 1,\; x \geq 0 \right\}
\end{aligned}
$$

and refer to Ax for $x \in \Delta$ as a *convex combination* of the columns of A. The convex hull of a set of points (or vectors) is the set of points "between" those vectors. We illustrate this here for $A \in \mathbb{R}^{2\times 2}$

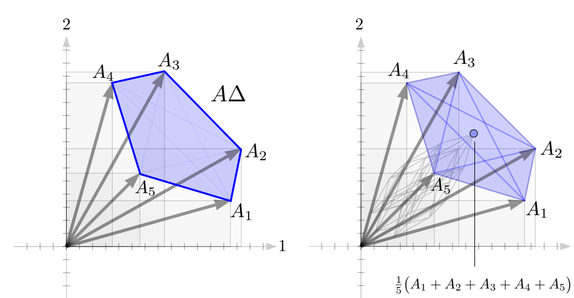


and $A \in \mathbb{R}^{3\times 3}$



Unlike the images of other types of sets, convex hulls do not actually depend on the location of the origin. The convex hull of any subset of columns is also within the convex hull (as is not hard to show). We also note that the arithmetic mean of the vectors (the average) is in the center of the convex hull.

For non-square A, the convex hull may be flat or the action of A may collapse dimensions of Δ. Here we show this set for $A \in \mathbb{R}^{2\times 5}$ with the outline of the convex hull illustrated. This case where A is fat is quite common in applications and often the outline of the convex hull (as opposed to the internal structure of the collapsed simplex is what is of interest. The projection of the edges and faces down onto the lower dimensional space may overlap significantly. The average of the points will still be near the center of the convex hull though more points to one side will naturally shift it toward those points. We also illustrate the average here with the hypercube defined by the coordinates $\frac{1}{5}\mathbf{1}$ illustrated.



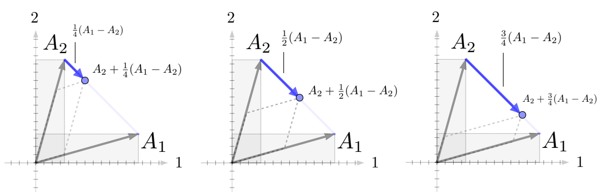
For matrices $A^{2\times 2}$ there is a specific parametrizetion of the convex hull that is often useful. In this case the set is a line segment between the two vectors and can be parametrized by a single variable α. The constraints on $x \in \mathbb{R}^2$ can be rewritten

$$\left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 \;\middle|\; x_1 + x_2 = 1,\ x_1, x_2 \geq 0 \right\}$$
$$= \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \alpha \\ 1-\alpha \end{bmatrix} \in \mathbb{R}^2 \;\middle|\; 0 \leq \alpha \leq 1 \right\}$$

Here the convex hull is parametrized by starting at one vector (say A_2) when $\alpha = 0$ and proceeding along the difference between the two vectors as α is increased until reaching the other vector (A_1) when $\alpha = 1$.

$$\begin{aligned} x_1 A_1 + x_2 A_2 &= \alpha A_1 + (1-\alpha) A_2 \\ &= A_2 + \alpha (A_1 - A_2) \end{aligned}$$

This construction is illustrated here.



Similar constructions are possible for more than two columns where the set is parametrized as starting at one column and proceeding to the other columns while varying $n-1$ parameters $\{z_j\}_{j=1}^{n-1}$. We list one such construction here that may be of use. Consider the matrix consisting of the differences between each column and a specific column (we choose the first column)

$$\begin{bmatrix} | & & | \\ A_2 - A_1 & \cdots & A_n - A_1 \\ | & & | \end{bmatrix} = AW$$

Note here that this matrix is given by AW where $W \in \mathbb{R}^{n \times n-1}$ is given by

$$W = \begin{bmatrix} - \ \mathbf{1}^T \ - \\ I \end{bmatrix} = \begin{bmatrix} -1 & \cdots & -1 \\ 1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 1 \end{bmatrix}$$

The convex hull of A can be denoted

$$\begin{aligned} A\Delta = \{ y \in \mathbb{R}^m \mid y = A_1 + AWz, \\ \mathbf{1}^T z \leq 1,\ z \geq 0,\ z \in \mathbb{R}^{n-1} \} \end{aligned}$$

This construction is illustrated here for $A \in \mathbb{R}^{3 \times 4}$

It is worth noting the relationship between these two characterizations of convex sets. For a point $y \in A\Delta$ we have that

$$y = A_1 + AWz = AI_1 + AWz = A(I_1 + Wz)$$

Here we have that x

$$x = I_1 + Wz = \begin{bmatrix} 1 - \mathbf{1}^T z & z_1 & \cdots & z_{n-1} \end{bmatrix}^T$$

Note here that if $z_j \geq 0$ and $\mathbf{1}^T z \leq 1$ then each $x_i \geq 0$. Note also that

$$\mathbf{1}^T x = \mathbf{1}^T (I_1 + Wz) = \mathbf{1}^T I_1 + \mathbf{1}^T W z = \mathbf{1}^T I_1 = 1$$

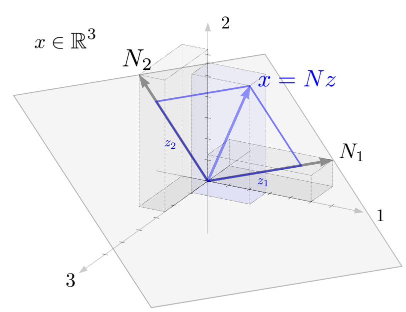
Remark 5. *In general, the above constructions are more useful for smaller x. If $x \in \mathbb{R}^n$, reducing to $z \in \mathbb{R}^{n-1}$ usually does not give much advantage. If $x \in \mathbb{R}^3$, reducing to $x \in \mathbb{R}^2$ can often be useful; and if $x \in \mathbb{R}^2$ (as in the first example), reducing to $z \in \mathbb{R}$ actually turns a vector problem into a scalar problem which can be very useful.*

5) Subspace Representations: We now turn to discussing images of subspaces.

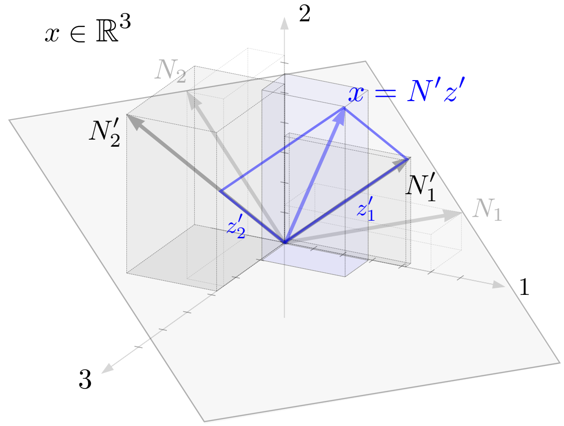
The first subspace representation

$$\mathcal{S}_1 = \{ x \in \mathbb{R}^n \mid x = Nz,\ z \in \mathbb{R}^k \}$$

is naturally thought of as the range of N.



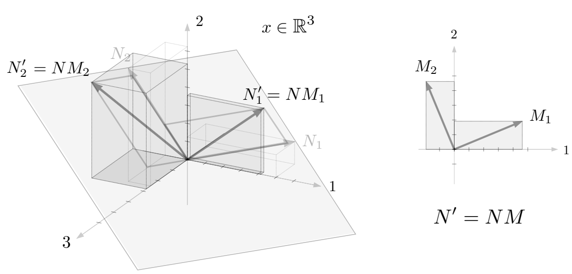
Note that for a particular subspace, the choice of N is not unique; it is only important that they have the same span. For a particular 2D subspace in $\mathbb{R}^3$, we illustrate two matrices N and N' that have the same span be used to represent the same subspace.



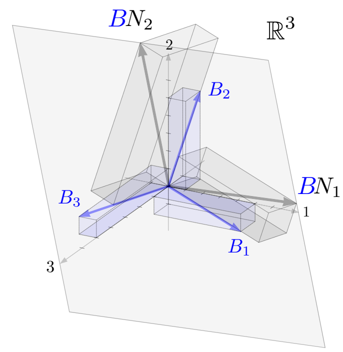
It is worth noting the relationship between different representations N and N'. Specifically there exists a matrix $M \in \mathbb{R}^{k \times k'}$ such that $N' = NM$ and

$$x = N'z' = NMz' = Nz$$

and thus we have the relationships $N' = NM$ and $z = Mz'$. We illustrate this here.



6) Subspace Images: Visualizing the image of domain subspaces when they are mapped into the co-domain is straight forward in that we just simply see them relative to the columns of A. Using the representations above, we can simply apply the matrix transformation to each column of N. We illustrate this here for the matrix $B \in \mathbb{R}^{3\times 3}$ shown.



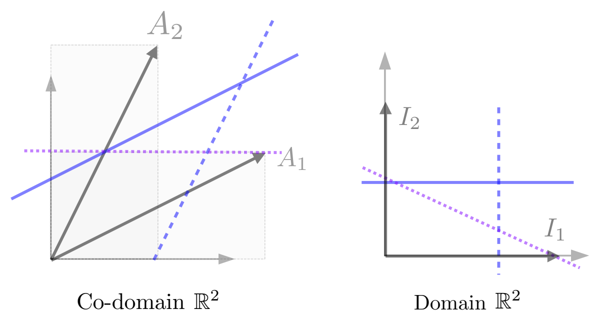
It is also worth giving several simple examples. First, the image of basic subspaces defined by fixing one coordinate, ie. of the form

$$\{x \mid x_j = \text{const}_j\}$$

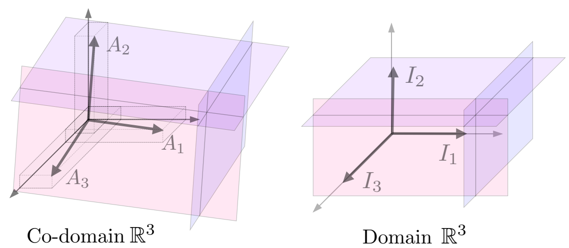
These affine spaces critical in that they are orthogonal to each of the coordinate axes. We illustrate them in for $A \in \mathbb{R}^{2\times 2}$ Another set of subspaces that particularly useful are the ones that form the faces of the 1-norm ball in the domain. These sets have the form

$$\{x \mid \bar{\mathbf{1}}^T x = 1\}$$

where $\left[\bar{\mathbf{1}}^T\right]_j = \pm 1$ is a signed summation vector. The image of these subspaces are particularly easy to visualize relative to the columns as shown in the figure below.



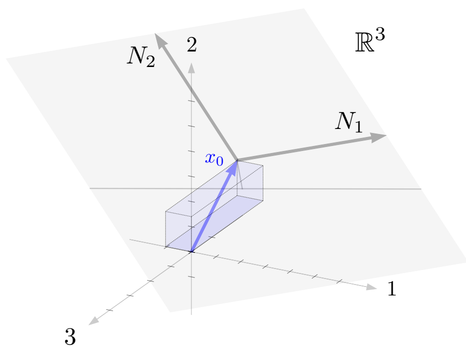
Note carefully in these examples how the direction of a particular column A_j actually has no impact on the subspace orthogonal to the jth coordinate vector and actually it is the other columns $\{A_i\}_{i \neq j}$ whose span determines these subspaces. This is immediate from the fact that $x_j = 0$ and thus the image $Ax = A_1x_1 + \cdots + A_nx_n$ should be independent from A_j these points x; but the affect on the geometry is worth noting in the domain.



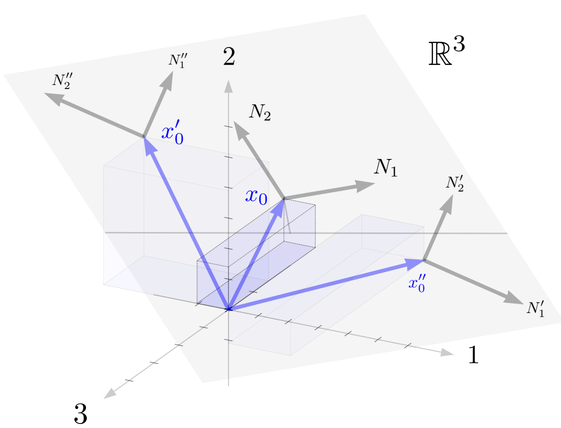
7) Affine Spaces: The first characterization of affine spaces

$$\mathcal{A}_1 = \left\{ x \in \mathbb{R}^n \;\middle|\; x = Nz + x_0,\ z \in \mathbb{R}^k \right\}$$

is simply a subspace shifted by x_0. We illustrate this here.



Similar to the discussion for subspaces above, for a particular affine space the choice of both N and x_0 are not unique. We illustrate three possibilities for the affine space shown above.



In fact, x_0 can be any point in the space and the columns of N can be any basis for the appropriate subspace. (Actually the columns of N need not even be linearly independent, though if they aren't than there will be redundant coordinates z that reach the same point in the space.) It is worth noting algebraically the relationship between these different representations. For two separate representations (N, x_0) and (N', x_0'), we have $N' = NM$. Note that the difference $x_0' - x_0$ is in the subspace and thus there exist coordinates such that $Nu = x_0' - x_0$. For any point x in the affine space

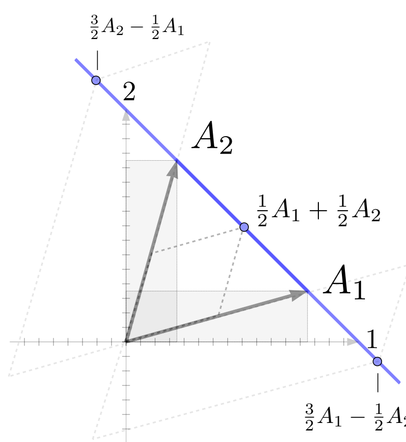
$$\begin{aligned} x &= N'z' + x_0' \\ &= NMz' + x_0' \\ &= NMz' + x_0' - x_0 + x_0 \\ &= NMz' + Nu + x_0 \\ &= N(Mz' + u) + x_0 \end{aligned}$$

Thus we have the relationships $N' = NM$ and $z = Mz' + u$

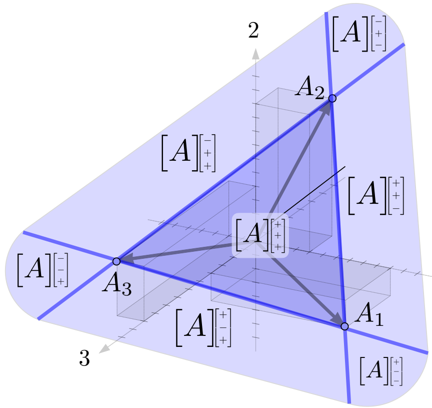
The third affine space representation

$$\mathcal{A}_3 = \left\{ x \in \mathbb{R}^n \;\middle|\; x = Az,\ \mathbf{1}^T z = 1, z \in \mathbb{R}^{k+1} \right\}$$

is an extension of the idea of a convex hull. Here the positivity constraints on each z_i are removed. Here the columns of A define k points that the subspace must pass through. Note that the extra constraint $\mathbf{1}^T z = 1$ reduces the dimension of the set by 1. In an m-dimensional space, k (with $k \leq m$) linearly independent points naturally defines an $k-1$ dimensional subspace that passes through all of them. For example, in $\mathbb{R}^2$, two points define a line; in $\mathbb{R}^3$ three points define a plane; etc. The image of this set through a matrix $A \in \mathbb{R}^{n \times k}$ is the $k-1$-dimesional affine space passing through all the columns. This is illustrated here for $A \in \mathbb{R}^{2\times 2}$ with several points labeled

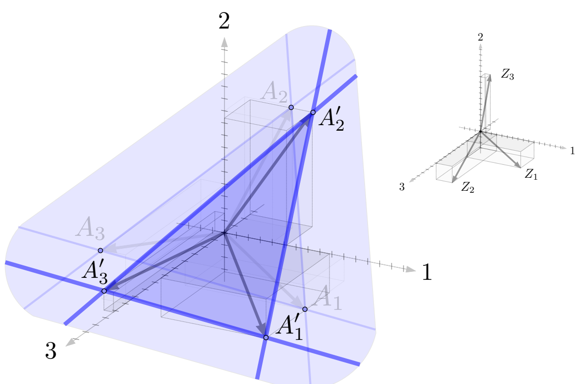


and for $A \in \mathbb{R}^{3\times 3}$ with the regions defined by z's with various sign profiles (though all summing to 1)



It is fairly straightforward to see how these sets extend the convex hulls of the points. These constructions can be quite useful when parametrizing lines and planes in $\mathbb{R}^2$ and $\mathbb{R}^3$.

Again for a given affine space, the representation is not unique. We illustrate two column representations A and A' here.



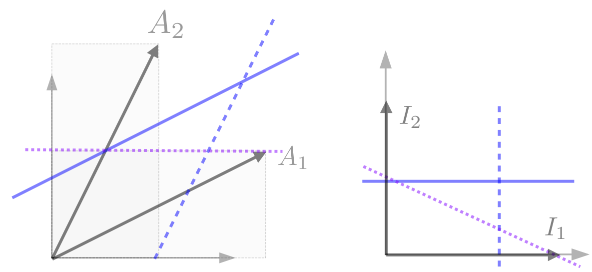
Given two representations of the space based on the columns of A and A', we can derive a relationship between the two representations. Since the columns of A' are in the affine space, we can write $A' = AZ$ where the columns of Z sum to 1, ie. $\mathbf{1}^T Z = \mathbf{1}^T$. Given this construction we have that

$$x = A'z' = AZz'$$

with $z = Zz'$. Note that $\mathbf{1}^T z = \mathbf{1}^T Zz' = \mathbf{1}^T z' = 1$. Here we can actually see A' as the columns of Z transformed by A as illustrated here.

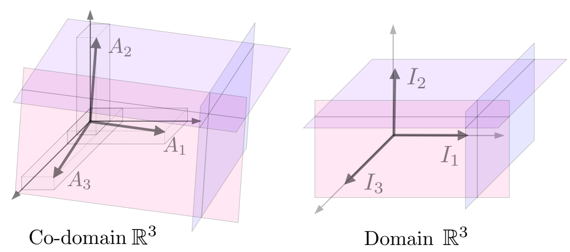
We also note several specific examples that are useful for practice and intuition. First suppose $B = \beta I$. In this case we are simplying scaling the length of the columns of A as shown here. This has the same effect as as modifying the constraint on x from $\mathbf{1}^T x = 1$ to $\mathbf{1}^T x = \beta$. Another example is letting B be a diagonal matrix $B = \mathrm{dg}(b)$ as shown here. Here we are scaling the columns of A each by different lengths. If the elements of b are ± 1, the image becomes alternative faces of the 1-norm ball. Next, we show examples where B is some subset of columns of the identity, specifically $B = \begin{bmatrix} I_1 & I_2 & I_3 \end{bmatrix}$ and $B = \begin{bmatrix} I_1 & I_2 \end{bmatrix}$ for $A \in \mathbb{R}^{2\times 4}$. Many other examples are obviously possible.

8) Affine Space Images: Visualizing the image of domain affine spaces when they are mapped into the co-domain is straight forward in that we just simply see them relative to the columns of A. We will be more precise shortly, but first we just give several basic examples for square $A \in \mathbb{R}^{2\times 2}$ and $A \in \mathbb{R}^{3\times 3}$. The fundamental visualization idea here is to "see" the domain relative to the columns of the matrix A.



Co-domain $\mathbb{R}^2$ Domain $\mathbb{R}^2$

Note carefully in these examples how the direction of a particular column A_j actually has no impact on the subspace orthogonal to the jth coordinate vector and actually it is the other columns $\{A_i\}_{i\neq j}$ whose span determines these subspaces. This is immediate from the fact that $x_j = 0$ and thus the image $Ax = A_1x_1 + \cdots + A_nx_n$ should be independent from A_j these points x; but the affect on the geometry is worth noting in the domain.



We show this here for several specific useful subspaces. We note that this is by no means an exhaustive presentation and we suggest that the reader attempt using this method for other subspaces of interest. We first show the image of basic subspaces defined by fixing one coordinate, ie. of the form

$$\{x \mid x_j = \text{const}_j\}$$

These affine spaces critical in that they are orthogonal to each of the coordinate axes. We illustrate them in for $A \in \mathbb{R}^{2\times 2}$ Another set of subspaces that particularly useful are the ones that form the faces of the 1-norm ball in the domain. These sets have the form

$$\{x \mid \bar{\mathbf{1}}^T x = 1\}$$

where $\left[\bar{\mathbf{1}}^T\right]_j = \pm 1$ is a signed summation vector. The image of these subspaces are particularly easy to visualize relative to the columns as shown in the figure below. We note also that one general way to define affine spaces is the set of points that satisfy

To be more precise,

Visualizing images of general affine spaces represented in this way, again, is simply a matter of transforming the columns of A. We illustrate this here for a matrix $B \in \mathbb{R}^{3\times 3}$ for affine space defined is similar to visualizing images of shapes and subspace. If we can "see" where each point in the domain maps in the co-domain we can apply this to each point in the affine space. To be precise, for visualizing images of affine spaces defined above through a matrix B we can simply apply the transformation B to the columns of A or the columns of N and x_0 and see how the space transforms. This is illustrated here briefly.

C. Domain Sets

Visualizing sets in the domain via column geometry is more subtle and difficult then visualizing sets in the co-domain but in many ways more fruitful.

We will start with a basic discussion of visualizing pre-images for matrices with full-column rank. Our first major payoff will come from a clean way to visualize the columns of A^{-1} (for invertible matrices) as the pre-image of the identity in the co-domain.

The first major subtlety will come with visualizing nullspaces of matrices that are column rank deficient (matrice that have non-trivial nullspaces). The matrices collapse directions (dimensions) of the domain and we will be concerned with visualizing these directions explicitly. We will spend a large portion of this next section on visualizing these nullspaces and nullspace representations of subspaces and affine spaces. Beyond just being geometrically interesting, these constructions will inform several algebraic techniques that are foundational for linear proofs. In particular, we will give very natural ways to construct nullspace bases and prove the rank-nullity theorem.

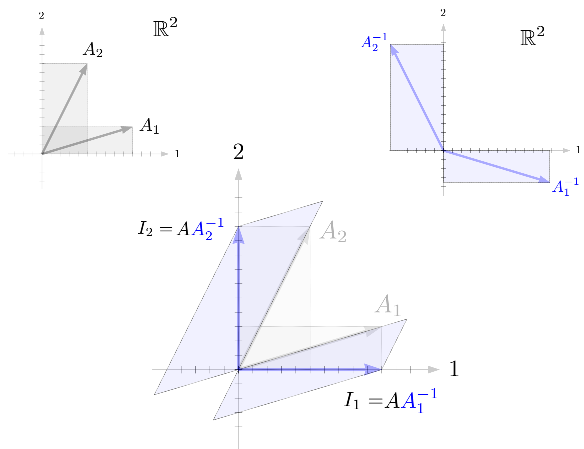
The second major subtlety in visualizing domain sets will come when we turn to visualizing the rows of the matrix in terms of the column geometry. Many sets of interest in the domain are directly related to the rows of A , most obviously the range of A^T. Visualizing the geometry of a matrix's rows in terms of it's columns is a surprisingly tricky exercise and perhaps the most unnatural we will discuss in this paper. We hope the readers will find our treatment satisfying.

Remark 6. *Visualizing the rows of a matrix A is the same as visualizing the columns of A^T. Algebraically, computing a matrix inverse is far more complicated than computing a matrix transpose. Visually, however, we will see that this is directly reversed. For an invertible A, visualizing the columns of A^{-1} will be much more natural than visualizing the columns of A^T. It is unclear if this is an insightful remark.*

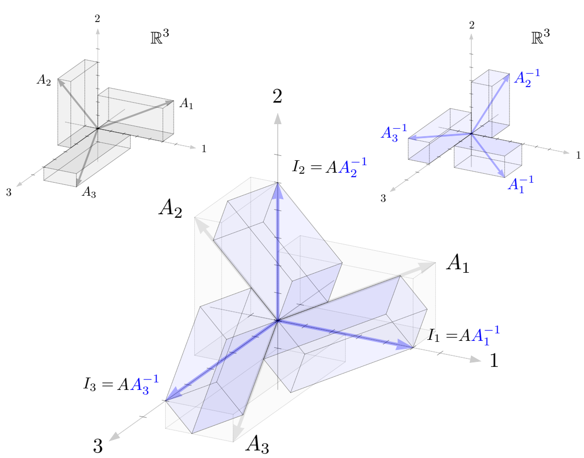
1) Basic Pre-Images: It is a bit more subtle to visualize the pre-image of co-domain sets before they are transformed through the matrix. This is done by drawing the set in the co-domain and then visualizing which points in the domain would end up in that set. Several examples for $A \in \mathbb{R}^{2\times 2}$ are illustrated here.

Note that opposite the image case discussed above, a small column tends to expand the relative size of a set in the preimage and a large column tends to shrink it. This is due to the fact that the columns represent the transformed standard basis vectors. If a standard basis vector will grow substantially when transformed by A, a small component in that direction in the pre-image can produce a large component in the image. The non-square (low-rank) case for pre-images is even more subtle than for images. For a tall (column rank deficient) matrix, there will likely be points in the image that are not within the range of A and thus have no corresponding points in the pre-image. For a fat (row rank deficient) matrix, any point in the nullspace of A will be deleted by A and is thus in the pre-image. Thus pre-images of sets for fat matrices are generally not compact. We will revisit this in our discussion of nullspaces below.

2) Inverses: One of the surprisingly straight-forward applications of visualizing a pre-image is visualizing the columns of a matrix inverse. The columns of A^{-1} are simply the pre-image of the identity I. We illustrate this here for several matrices $\mathbb{R}^{2\times 2}$



and $\mathbb{R}^{3\times 3}$



This technique is surprisingly simple and should allow visually oriented readers to estimate inverses for 2×2 and 3×3 matrices quite quickly.

Remark 7. *There is a dynamic version of this visualization that can be fruitful. A reader with n-hands could picture grabbing each column of A and pulling them to the appropriate standard basis vector. As the columns move, the rest of the space gets dragged/stretched with them. As the columns of A move to the standard basis vectors, the standard basis vectors will move to the columns of A^{-1}. (Try this mentally for the $2x2$ example shown above.)*

3) Nullspace of A: We now turn to subspaces in the domain starting with the nullspace. We note that in this discussion it is useful to keep any intuition preimage intuition from the above discussion in mind. The nullspace of A is the preimage of the point 0. Another way to say this is the coordinates of the point 0. Note that in order for a non-trivial nullspace to exist there must be more columns in A than dimension of the span of A in the co-domain. There is a classic construction for constructing a basis for a nullspace that will aid our visualizations. For a matrix $A \in \mathbb{R}^{m\times n}$ with rank k, 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

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

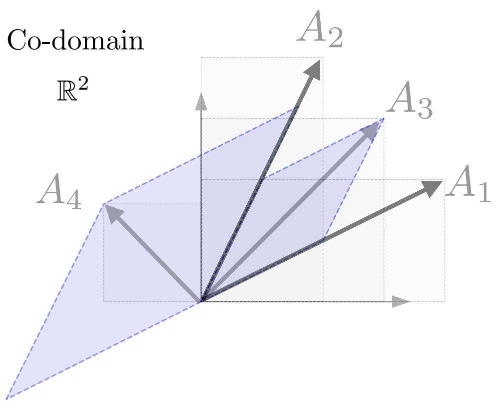
then $AN = 0$. With a little more work, we can show that the columns of N form a basis for the nullspace. The identity block proves the linear independence of the columns. We can also show explicitly that any element in the nullspace of A is in the span of N. This fact relies on the linear independence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

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

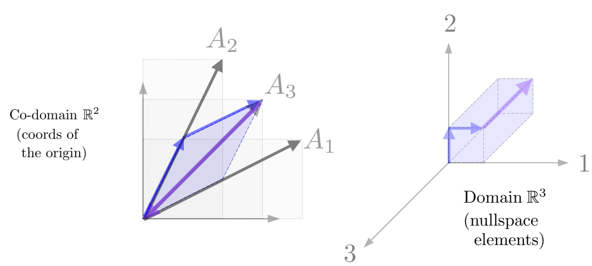
$$\text{SPAN:} \qquad Ax = 0 \Rightarrow A'\begin{bmatrix} I & B \end{bmatrix}\begin{bmatrix} x' \\ x'' \end{bmatrix} = 0$$

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

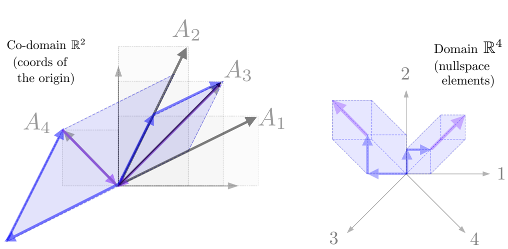
where the beginning the second line depends on the columns of A' being linearly independent. Writing the columns of A'' in terms of the columns of A' is illustrated in the figure XXX below.



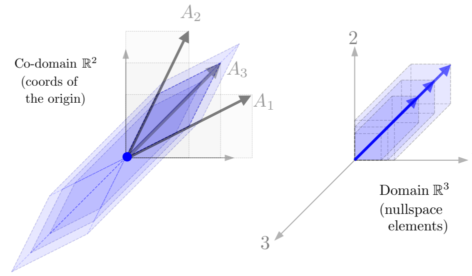
We note in this case, we are actually visualizing specific (non-zero) points in the domain that map to 0. If we expand these out in the domain we get the points in these figure illustrated for a matrix with one "extra" column $A \in \mathbb{R}^{2\times 3}$



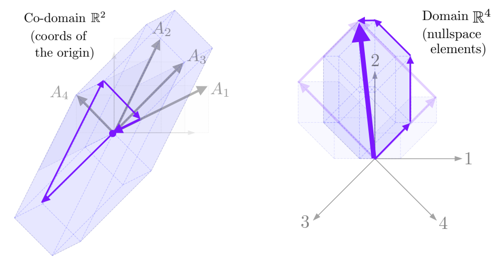
and for a matrix with two extra columns $A \in \mathbb{R}^{2\times 4}$.



Any linear combination of these vectors is also in the nullspace of A and thus the span of these vectors is the nullspace as illustrated for $A \in \mathbb{R}^{2\times 3}$ (a one dimensional nullspace)



and for $A \in \mathbb{R}^{2\times 4}$ (a richer two dimensional nullspace example)



Note that in both cases, the element in the nullspace provide a way to move away from the origin and then back to it along the directions given by the columns of A.

If one wishes, one could visualize pulling the columns of A back to the axes of the domain and watching the point 0 expand (possibly in multiple directions) to points in the nullspace. There is a variation of this expansion idea that can help us visualize nullspace but it is more natural in the context of visualizing affine spaces below.

The nullspace constructions given above are critical for defining subspaces in general. Specifically, there are two natural ways to define a dimension k subspace of $\mathbb{R}^n$: one as the nullspace of a rank k matrix $A \in \mathbb{R}^{m\times n}$ and one as the range of a matrix $N \in \mathbb{R}^{n\times k}$.

$$\mathcal{X} = \left\{ x \in \mathbb{R}^n \;\middle|\; Ax = 0 \right\}$$
$$\mathcal{X} = \left\{ x \in \mathbb{R}^n \;\middle|\; x = Nz,\;\; z \in \mathbb{R}^k \right\}$$

For the two characterizations to be of the same subspace, we want to choose N as above to be a basis for the nullspace of A.

D. *Affine Spaces (Domain)*

The nullspace discussion above extends to visualizing sets of the form

$$\mathcal{X} = \left\{ x \in \mathbb{R}^n \;\middle|\; Ax = b \right\}$$

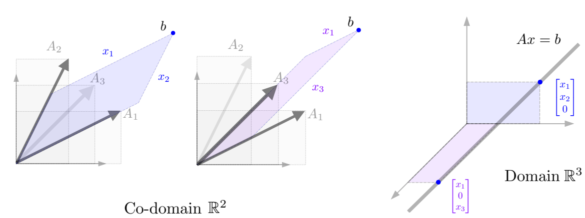
for $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and $k < n$. These type of sets can be more explicitly characterized using a basis for the nullspace of A written as the columns of the matrix $N \in \mathbb{R}^{n\times k}$ using the form

$$\mathcal{X} = \left\{ x \in \mathbb{R}^n \;\middle|\; x = Nz + x_0,\ z \in \mathbb{R}^k \right\}$$

where x_0 is a specific solution to the equation $Ax = b$. Similar to the construction above, we can find coordinates of b relative to the columns of A' such that $b = A'u$ in order to determine a specific solution of the form $x_0^T = \begin{bmatrix} u^T & 0 \end{bmatrix}^T$. Note that other specific solutions could be used by using different columns of A as basis vectors or by using linear combinations of more than k vectors in the matrix. We illustrate these possible solutions for $A \in \mathbb{R}^{2\times 3}$. In general for a rank k matrix a basis for the co-domain requires k columns. Assuming any set of k columns is linearly independent, we can pair the first $k-1$ columns with any of the other $n-k+1$ columns to form a basis. It is not immediately obvious but specific solutions computed from bases of this form are sufficiently rich to span all specific solutions (see Appendix XXX for further explanation and full proof). For the 2×3 case given above, this reduces to two bases comprised of columns $\begin{bmatrix} A_1 & A_2 \end{bmatrix}$ and $\begin{bmatrix} A_1 & A_3 \end{bmatrix}$. If we compute specific solutions for b with these bases we get two possible solutions $x, x' \in \mathbb{R}^3$ of the form

$$x = \begin{bmatrix} x_1 & x_2 & 0 \end{bmatrix}^T, \qquad x' = \begin{bmatrix} x_1' & 0 & x_3' \end{bmatrix}^T$$

such that $Ax = b$ and $Ax' = b$. These solutions are illustrated in the figure below.



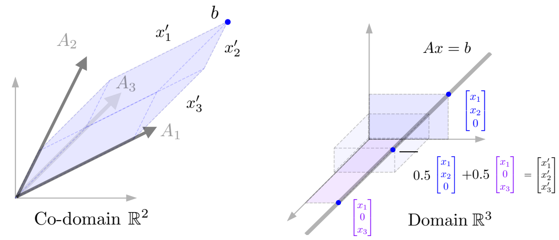
The line between these two points can be defined as the set of points

$$\left\{ \alpha x + \alpha' x' \;\middle|\; \alpha + \alpha' = 1 \right\}$$

Algebraically, we can see easily that these points are also solutions to the affine equation given above. Explicitly,

$$A(\alpha x + \alpha' x') = \alpha Ax + \alpha' Ax' = \alpha b + \alpha' b = b$$

We illustrate this here



Further more we can see that the difference between these two points (or any scalar multiple of this difference) is in the nullspace of A. Again explicitly

$$A(x - x') = Ax - Ax' = b - b = 0$$

Comparing this picture with the nullspaces images above illustrate this fact.

Thus the picture that we have is that these two solutions define a 1D subspace (a line) all of whose points satisfy the affine equation $Ax = b$. We also have that vectors along the line are in the nullspace of A. We now move to the case where $A \in \mathbb{R}^{2\times 4}$. Here, we can form three separate bases from the columns of $\begin{bmatrix} A_1 & A_2 \end{bmatrix}$, $\begin{bmatrix} A_1 & A_3 \end{bmatrix}$, and $\begin{bmatrix} A_1 & A_4 \end{bmatrix}$. The coordinates of b relative to these bases are illustrated in the following figure as well as the three specific solutions

$$x = \begin{bmatrix} x_1 & x_2 & 0 & 0 \end{bmatrix}^T,\ x' = \begin{bmatrix} x_1' & 0 & x_3' & 0 \end{bmatrix}^T,\ \ x'' = \begin{bmatrix} x_1'' & 0 & 0 & x_4'' \end{bmatrix}^T,$$

Again as in the case above, any combination of these solutions with coefficients summing to one is also a solution. The set

$$\left\{ \alpha x + \alpha' x' + \alpha'' x'' \;\middle|\; \alpha + \alpha' + \alpha'' = 1 \right\}$$

is the 2D affine space shown in the figure. Again any point in this set is also a solution to $Ax = b$.

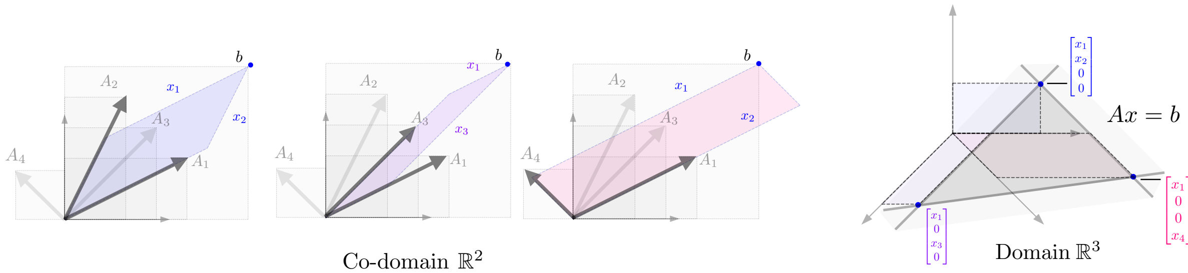
$$\begin{aligned} A\big(\alpha x + \alpha' x' + \alpha'' x''\big) &= \alpha Ax + \alpha' Ax' + \alpha'' Ax'' \\ &= (\alpha + \alpha' + \alpha'')b = b \end{aligned}$$

Pairwise differences between the solutions ($x - x'$, $x - x''$, etc.) can also be used to construct elements in the nullspace and a basis for the direction spanned in the affine space.

Several proofs for the general case are given in the appendix.

In general, the affine space given can be characterized by a specific solution x_0 plus any element in the nullspace. The solution x_0 can be chosen to be any of the elements given above. Often it is chosen to be the minimum norm solution, This is computed as

$$x_0 = A^T(AA^T)^{-1}b$$

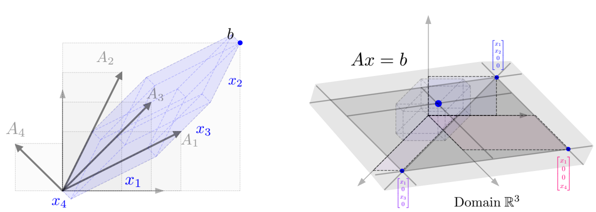


Nullspace column geometry example for $A \in \mathbb{R}^{2\times 4}$

This solution in general will use a linear combination of all the columns to construct b (all elements of x will be non-zero). Whereas our previous specific solutions were chosen by selecting a basis and then inverting that basis to find the specific solution from b, the minimum norm solution is chosen based on optimization principles to find the point on the affine space closest to the origin (as defined by the 2-norm), ie.

$$x_0 = \arg\min_x \|x\|_2^2 \quad \text{s.t.} \quad Ax = b$$

We illustrate this solution for $A \in \mathbb{R}^{2\times 4}$. Note the proximity to the origin of the point x_0 in the domain. And also the relative size of the hypercube representation of the coordinates in the co-domain picture.



The above constructions for affine spaces suggests several alternative method for visualizing nullspaces via column geometry. If we pick any point b and envision different sets of coordinates for b via the columns of A (say x and x' such that $Ax = b$ and $Ax' = b$), then the difference between these two points is a vector in the nullspace of A. The dimension of the nullspace is the number of truly independent ways to construct coordinates for b. Note also that this construction is independent of b and thus the choice of b does not matter. For an alternative more dynamic visualization, we could think of pulling each column of A back to an axis in the domain. As we pull, we start to see subspace of points that all expand out of the point b,ie. that would all map to the point b though A. These directions form the nullspace of A. If there is only one vector in the domain that maps to b then the nullspace is trivial (dimension 0).

The set of points is then the point x_0 plus any element in the nullspace. (Again we can envision watching the point b expand out to this affine set as we drag the columns of A back to the axes of $\mathbb{R}^n$ as shown in the inset.) We give several illustrations of this construction for different dimensions shown below. It is worth considering these examples quite carefully as well as visualizing the dimensions in the nullspace as degrees of freedom in choosing x_0. This visualization is conceptually complicated and worth noting how various aspects change as both b and the columns of A shift position and andthe position of the columns of A and as

Visual Exercise: Vary the following:

- Position of b
- Position of columns of A
- Number of columns of A (dim n)

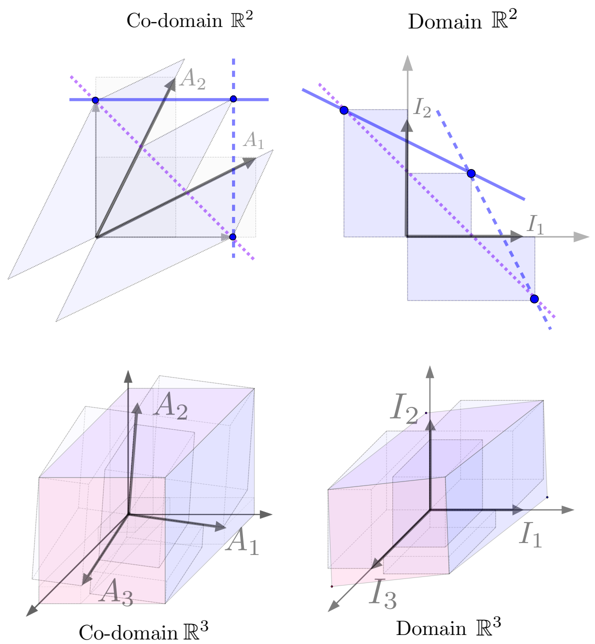
1) Pre-Image of Co-Domain Subspaces: While perhaps slightly more complicated, it is also very useful to visualize the pre-image of co-domain subspaces in the domain.

It is

It is also suWe focus our discussion on begin (and spend most of our effort) visualizing pre-images of sets where a coordinate value is fixed in the co-domain. Specifically,

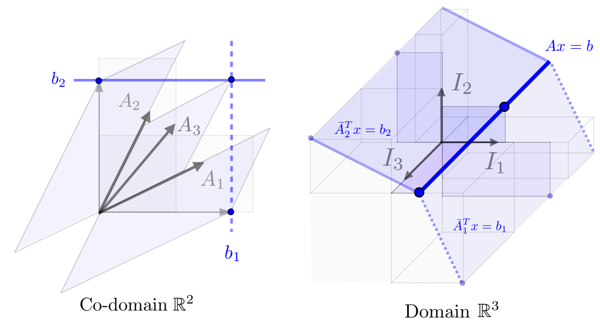
$$\{x \mid I_i^T Ax = b\} = \{x \mid \bar{A}_i^T x = b_i\}$$

In other words the set of points that satisfy one row of an affine constraint. We illustrate this for both rows



We now turn to the general case where A is not square, for tall A, part of the co-domain is unreachable through A and thus when we look at pre-images of co-domain sets, we must first intersect those sets with the range of A as illustrated in this figure. After we've done this intersection, our visualization proceeds as before.

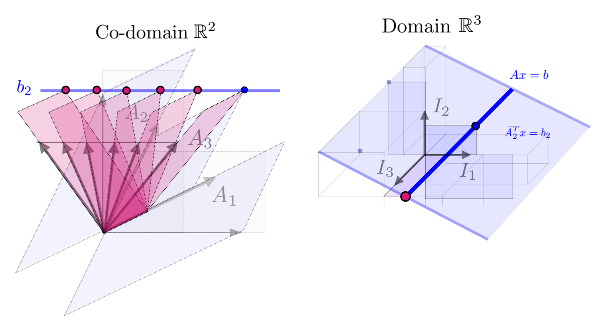
The fat case is more subtle and also much more relevant. In this case, A has more columns, ie. redundant directions, and thus the pre-image of any set also includes also any elements in the nullspace of A (as these elements will be deleted as they pass through A). For a rank k matrix A, the nullspace has rank $n-k$ (by rank-nulllity). Here some care should be given to make sure we see all appropriate dimensions in the pre-image. If our set in the co-domain is a simple point than the pre-image is an affine space with the dimension of the nullspace $n-k$ (as discussed at length in the section on affine spaces). If the the intersection of the set in the co-domain has dimension q, then the pre-image will in general have dimension $q+(n-k)$. Here, when we "see" the pre-image relative to the columns we almost visualize that pre-image having dimensions/degrees of freedom given by the nullspace.



Note that when we consider column geometry we are technically considering a vector space given by the number of constraints. From the above example, we note that we can also visualize individual constraints in this multi-dimesional. In the picture above, for example, if we ignore the constraint in the first dimension and just consider the constraint in the second direction than although we started with a 2×2 matrix, the second row of A defines the equation

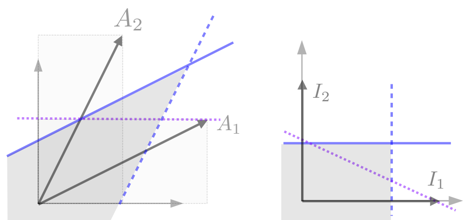
$$I_2^T Ax = I_2^T b = b_2$$

which is a single constraint on vectors in $\mathbb{R}^3$ and thus defines a $2D$ subspace pictured below. We note that from the equation above the first row does not have any affect on the constraint given. It is not at all obvious but if we shift any of the vectors, the columns of A or b in the horizontal dimension only, the picture of the domain does not change.



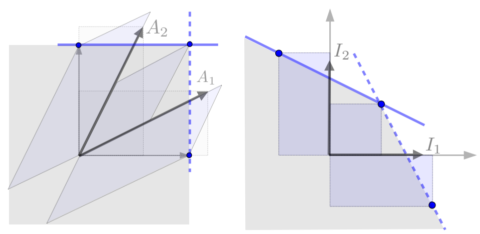
E. Inequality Column Geometry

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

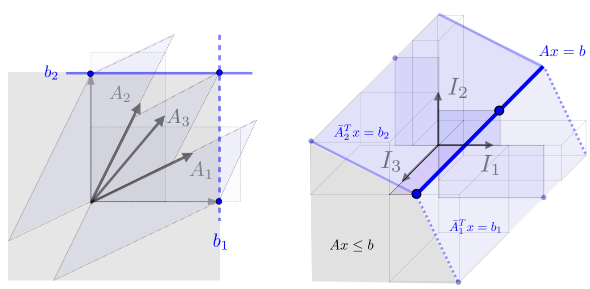


Co-domain $\mathbb{R}^2$ Domain $\mathbb{R}^2$

Co-domain $\mathbb{R}^2$ Domain $\mathbb{R}^2$



and preimage of inequality sets which are often more useful.



Co-domain $\mathbb{R}^2$ Domain $\mathbb{R}^3$

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

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

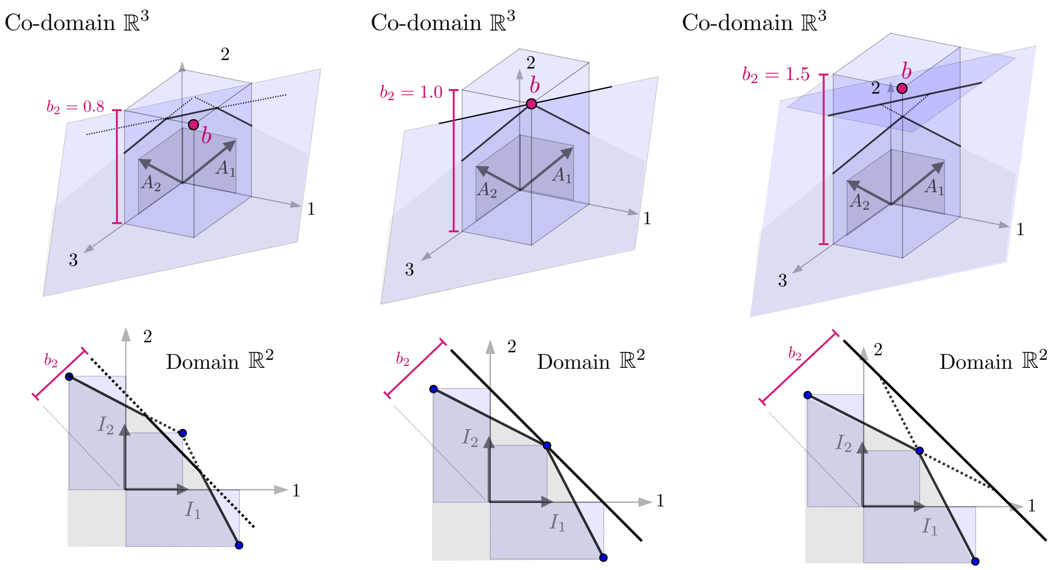
If we compute a basis for the nullspace of A and rewrite the affine constraint to have the form in it's nullspace form ($x = Nz + x_0$) then we can rewrite this as the image of a constraint on the variable z

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

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

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

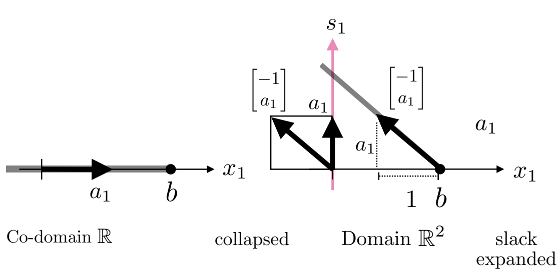
We now give several low dimensional examples of this construction that are illustrative. Note the dimensions of the matrices in each each case.



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.

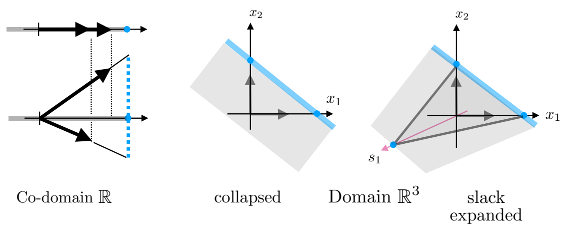


Note here that the dimension of the original inequality and the relaxed set are the same despite the fact that the second is in a larger ambient space. Note also the slack variable representatino of the set is the line $s = -c_1 x_1 + d_1$ in the (x_1, s_1) intersected with the positive half-space $s_1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

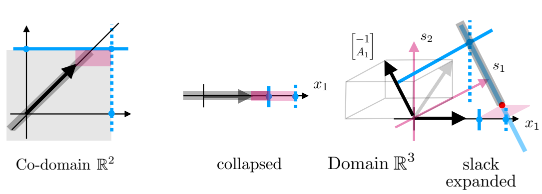
We now proceed to slightly richer examples. Specifically, we consider two of the most basic examples where C is not square. We start with the case where C is fat; $C = \begin{bmatrix}c_1 & c_2\end{bmatrix} \in \mathbb{R}^{1\times 2}$ and $d_1 \in \mathbb{R}$. In this case the constraints can be represented 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 $s_1 = d_1$.



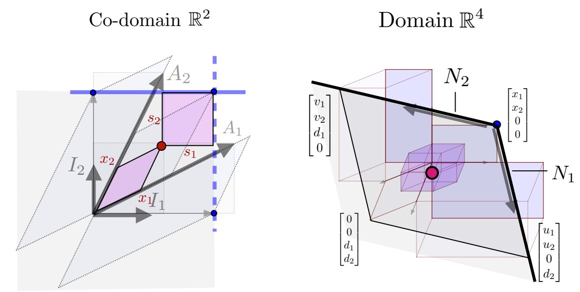
We now proceed to the case where C is tall; $C \in \mathbb{R}^{2\times 1}$. In this case $x = x_1$ is simply a 1D variable and the constraints provide two inequality constraints on the same variable. Often only one of them will be relevant. For example, take $c_1 = c_2 = 1$ and $d_1 = 1$ and $d_2 = 2$. 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.



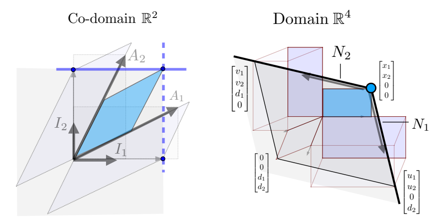
We lastly turn to the case where $C \in \mathbb{R}^{2\times 2}$

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

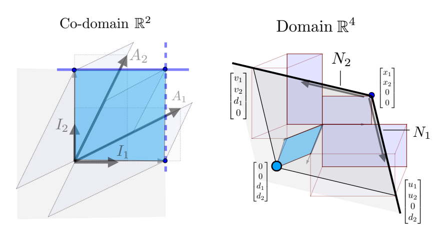
Here the original variables are 2D $x \in \mathbb{R}^2$ and two constraints correspond to two slack variables $s \in \mathbb{R}^2$. The slack variables lift the original inequality set to the intersection between the orthant $s \geq 0$ and the affine space defined by the matrix $\begin{bmatrix} C & I \end{bmatrix} \in \mathbb{R}^{2\times 4}$. 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.



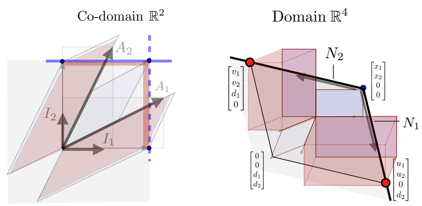
If $x = 0$, then the constraints must be satisfied only with slack variables and $s = d$.



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)

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

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 8. *These slack variable examples are subtle and*

VI. 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”.

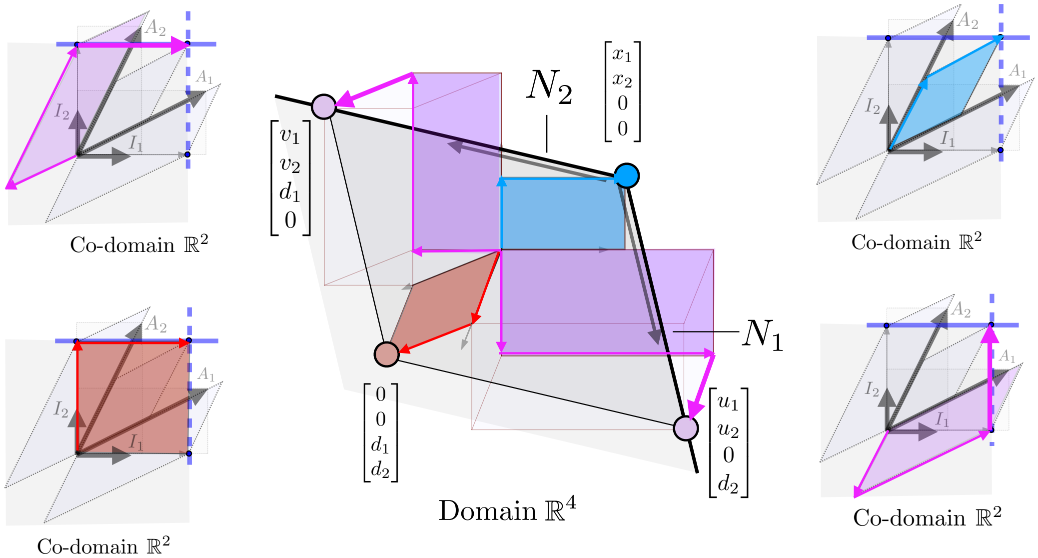
A. Set Transformations

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