Column Geometry Visualizations

Dan Calderone

May 28, 2023

Abstract

This tutorial paper gives basic visualizations for matrix column geometry. Matrix column geometry is presented first with a discussion of matrix-vector and matrix-matrix multiplication. The body of the paper is then divided into a discussion of co-domain sets and domain sets. In the co-domain, basic set images, the range and adjoint nullspace are discussed. Particularly focus is given to image representations of affine spaces. In the domain, basic pre-images, the nullspace and adjoint range space are discussed at length. Particularly focus is given to nullspace representations of affine spaces.

1 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.

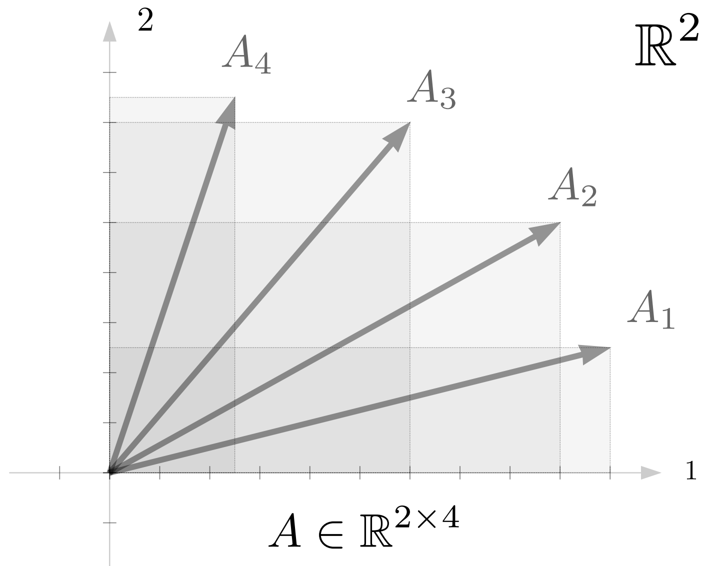
- Vector visualizations

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]

2 Basic Column Geometry

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

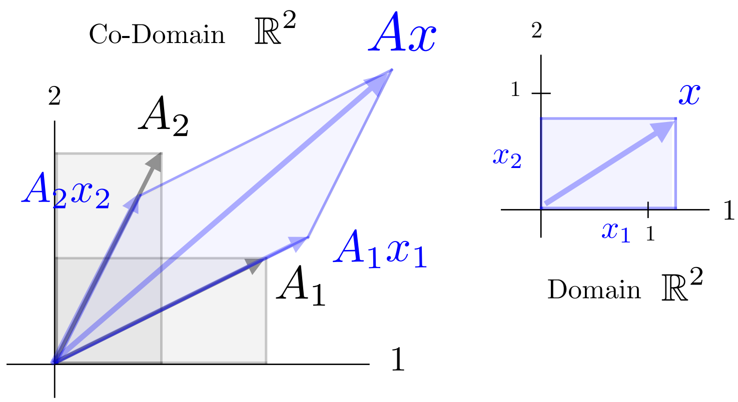
$$A = \begin{bmatrix} | & & | \\ A_1 & \cdots & A_n \\ | & & | \end{bmatrix}$$



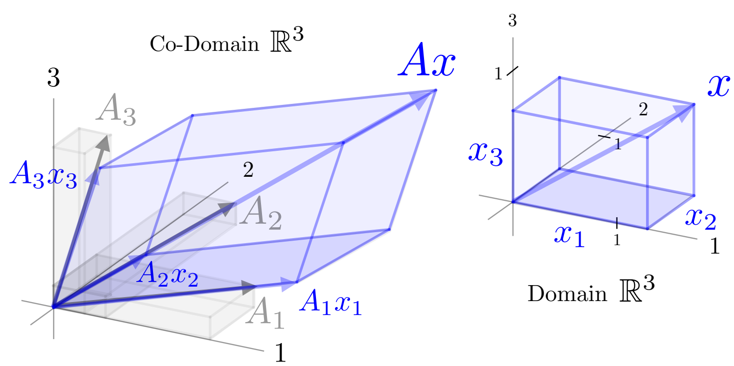
Each individual column $A_j \in \mathbb{R}^m$ tells where the jth standard basis vector (in the domain) gets mapped under the transformation. Explicitly $AI_j = A_j$. We can see where a vector $x \in \mathbb{R}^n$ in the domain gets mapped by breaking up x into a linear combination of standard basis 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 $A \in \mathbb{R}^{2\times 2}$

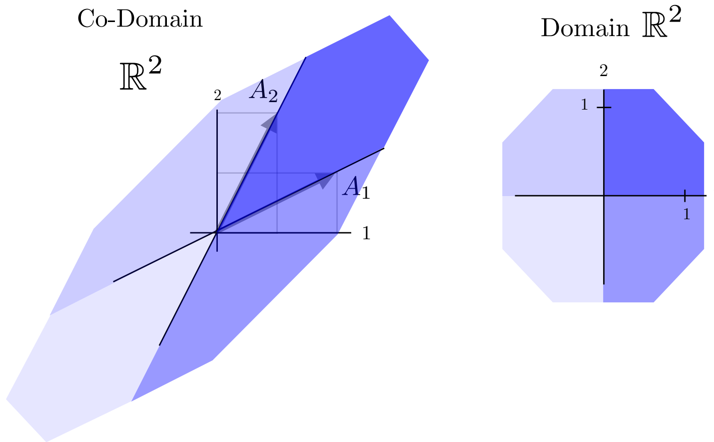


and also $A \in \mathbb{R}^{3\times 3}$.



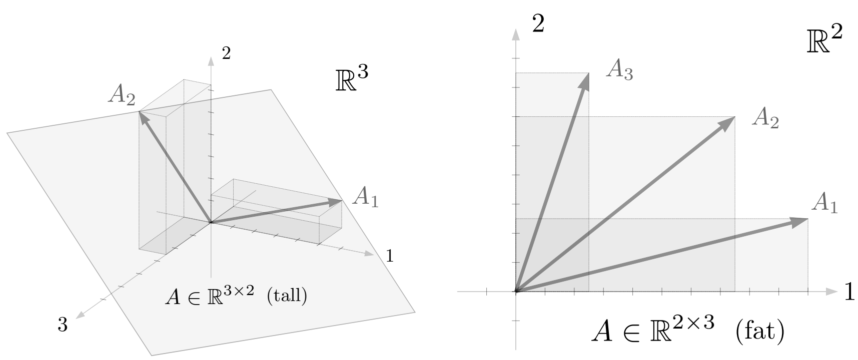
As in our discussion of vectors, the various routes from the origin to the tip of the vector along edges of the hypercube relate to the different order the scaled columns can be added in. Note also that if we take $A = I$, we simply get the vector itself back.

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.

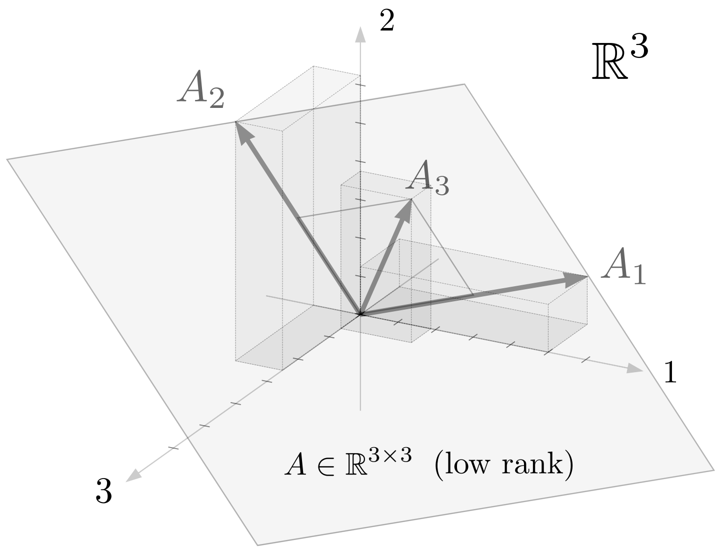


We will use this technique ad nauseam in what follows so it is worth getting comfortable with it.

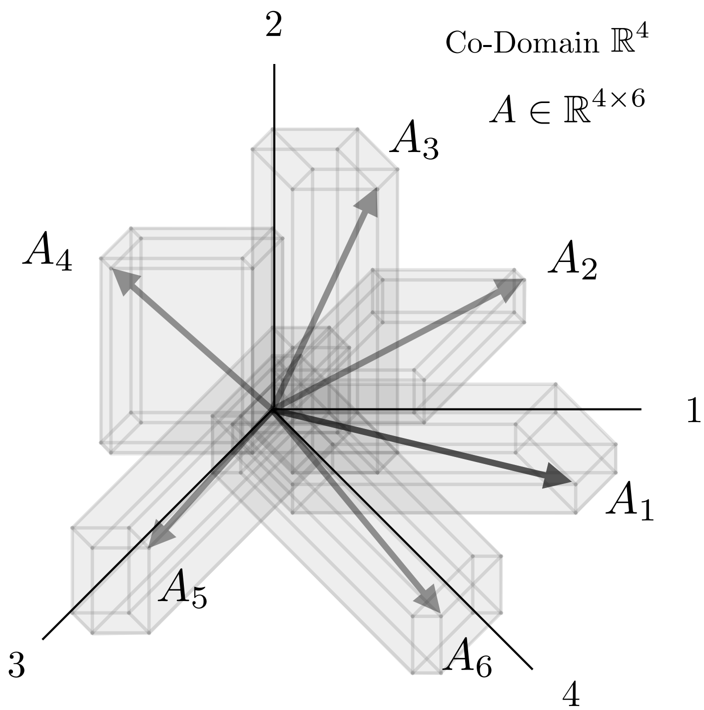
Note here there is nothing that immediately requires that A be square. For tall matrices (fewer columns than dimensions) the reachable points will be a subset of the co-domain; for fat matrices (more columns than dimensions) there will be redundant ways to reach any point in the co-domain. We give two 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.



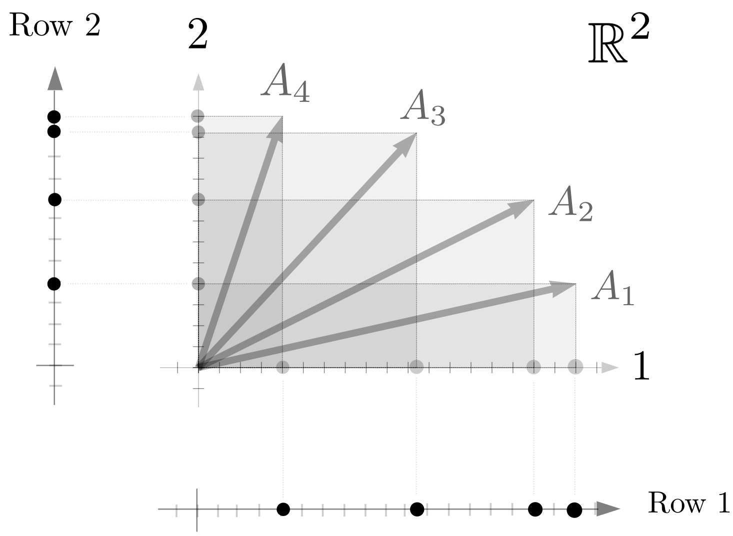
For more than two or three columns, the space we will be visualizing is high dimensional as in the vector illustrations in Figure XXX (with all the associated problems of depth). The co-domain of the map is ambient vector space that the columns live in and is thus easy to visualize. Depending on how 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.



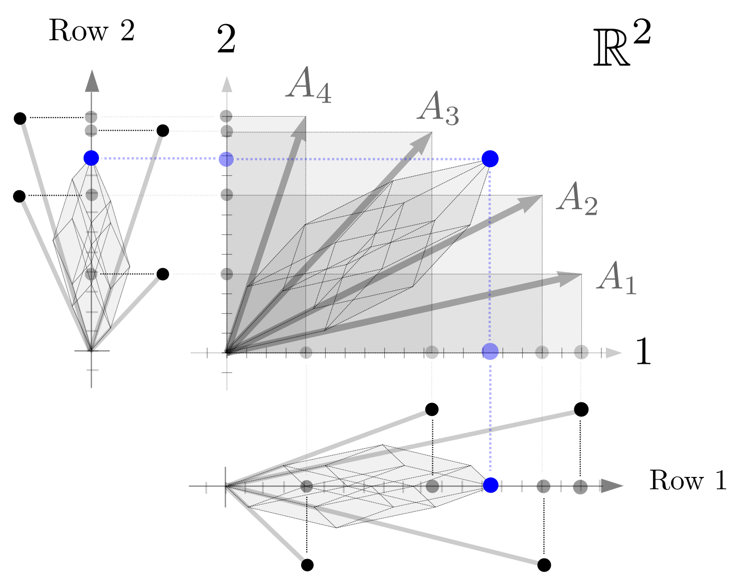
Remark 1. *If we want to focus in on a particular coordinate of the output, we project the output vector onto that particular axis. Algebraically, here we are considering the inner product between x and that particular row of A, $\bar{A}_j^T x$ We note that here if we just focus at the distance of the 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.*

2.1 Rows - Parallel Axis Representation

The most basic and in some ways natural way to "see" the rows of the a matrix via the columns is to note that the rows are represented in parallel on each axis in the space with elements given by the appropriate projection of each vector onto that axis as illustrated here.



The projection of an input vector onto each row of A can then be seen by using the hybrid parallel-spatial representation of the inner product.



Actually, we can note that the overall column geometry picture is just using the hybrid geometry inner product picture on both axes (rows) simultaneously. One visualization of the rows can be created by simply continuing the expansion process from the previous figure and lining up each coordinate with a new axis as visualized here.

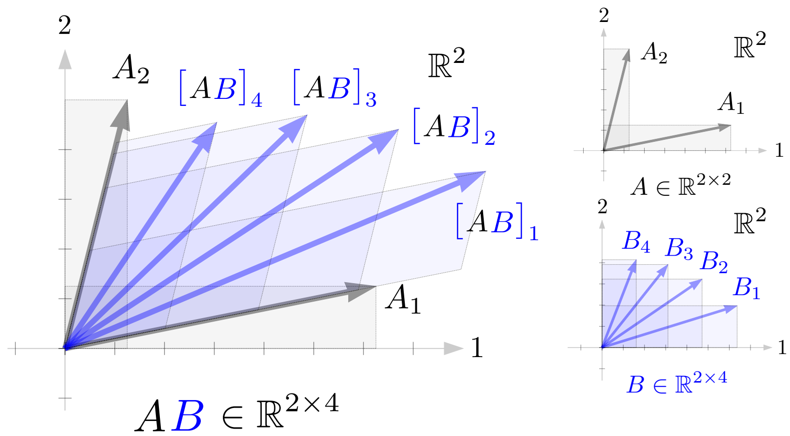
While accurate, this method involves perhaps too much squinting and motion.

2.2 Matrix Multiplication for Column Geometry

2.2.1 Right Multiplication

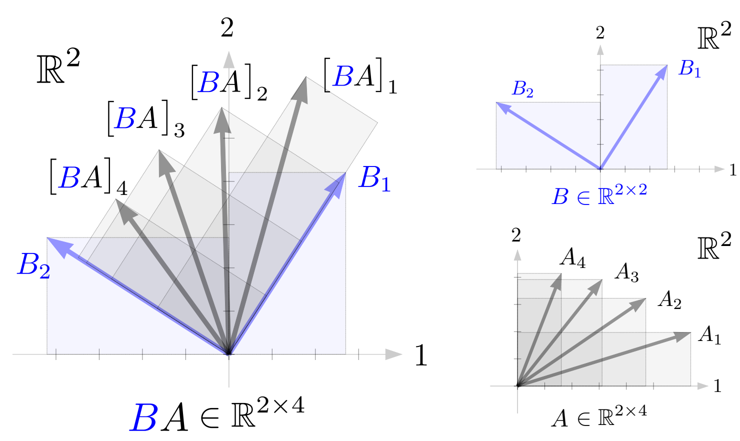
When we right multiply a matrix $A \in \mathbb{R}^{m \times n}$ by a matrix $B \in \mathbb{R}^{n \times p}$, each columns of B take linear combinations of the columns of A. Algebraically, this corresponds to treating each column of B as a separate vector that gets multiplied by A, ie.

$$AB = A\begin{bmatrix} B_1 & \cdots & B_p \end{bmatrix} = \begin{bmatrix} AB_1 & \cdots & AB_p \end{bmatrix}$$



2.2.2 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 2. *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.*

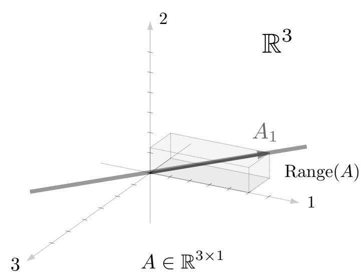
We now turn our focus to visualizing sets in the co-domain and the domain relative to the column geometry of a matrix. We start with the co-domain since column geometry provides a a natural way to visualize images of (domain) sets transformed by the matrix. The bulk of this section will focus on visualizing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

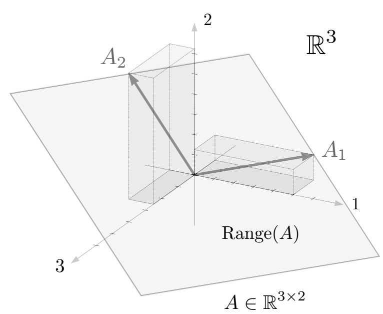
3 Co-Domain Sets

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.

3.1 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 $\mathbb{R}^3$, $A \in \mathbb{R}^{3\times 1}$ and $\mathbb{R}^{3\times 2}$.

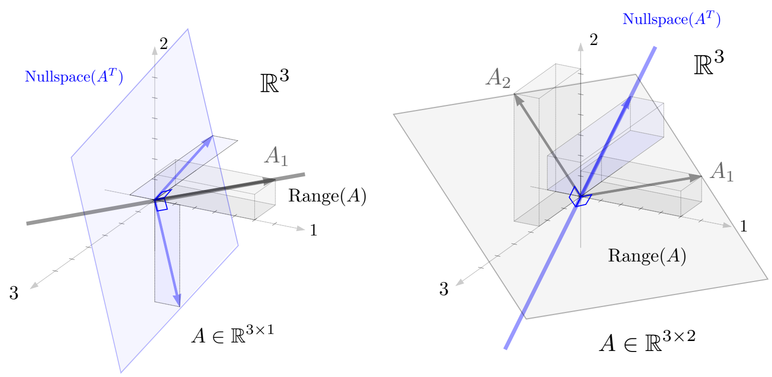




The discussion above gives several other good examples of matrix ranges.

3.2 Nullspace of A^T

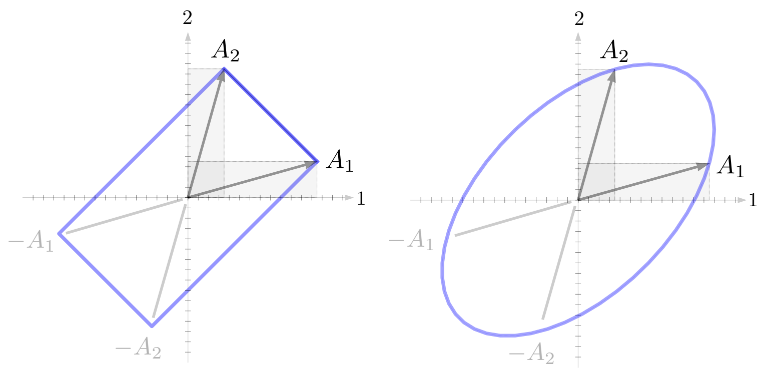
The orthogonal complement to the range of $A \in \mathbb{R}^{m\times n}$ is the set of vectors with no component in the range. This subspace corresponds with the nullspace of A^T, ie. is orthogonal to all the columns of A. We illustrate this here for the range examples above. Note that the dimension 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.



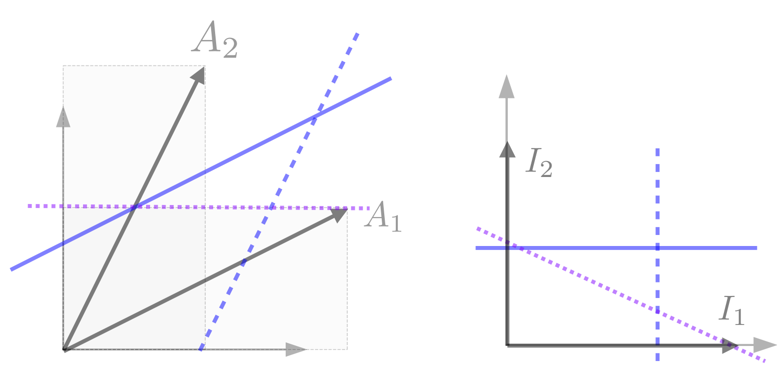
We now proceed to images of basic sets. Note that all the images that follow will be subsets of the range.

3.3 Basic Set Images

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



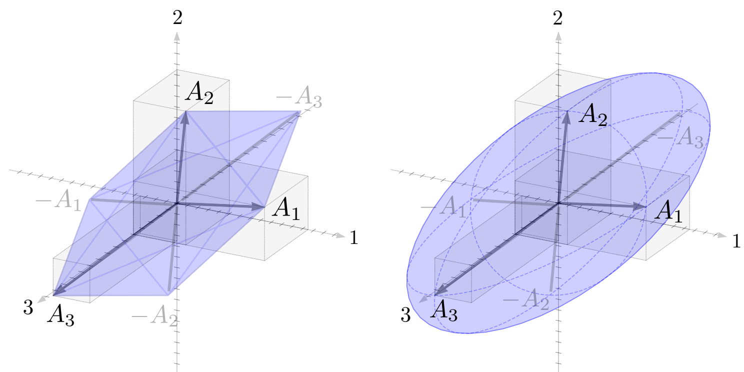
as well as images of several affine spaces.



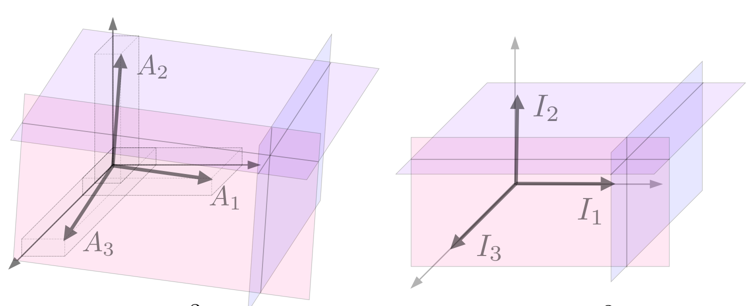
Co-domain $\mathbb{R}^2$ Domain $\mathbb{R}^2$

It is worth noting 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 also give several examples of the image of 1-norm and 2-norm through a matrix $A \in \mathbb{R}^{3\times 3}$

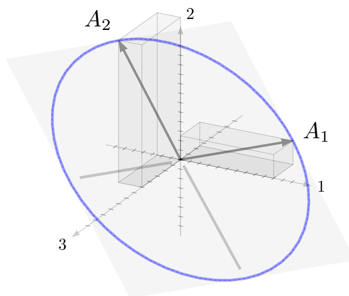


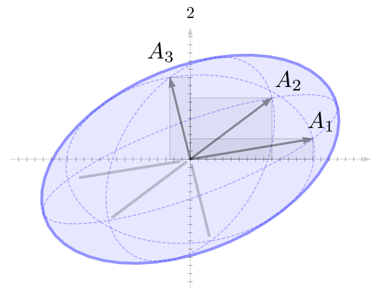
Note that the simplex is highlighted as well. These sets were chosen as they show the geomery of each orthant (in the domain) under the transformation. The location of each orthant should be noted by the reader. We also give examples of images of several affine spaces.



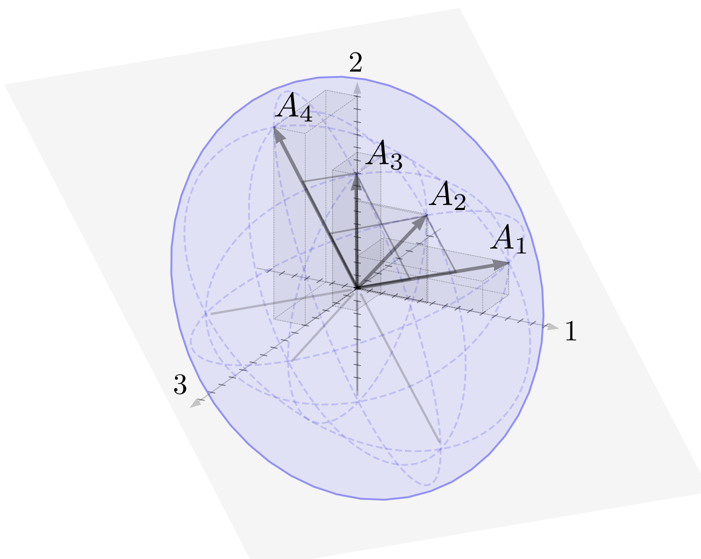
Co-domain $\mathbb{R}^3$ Domain $\mathbb{R}^3$

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 $A \in \mathbb{R}^{3\times 2}$ 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 $A \in \mathbb{R}^{2\times 3}$ 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.

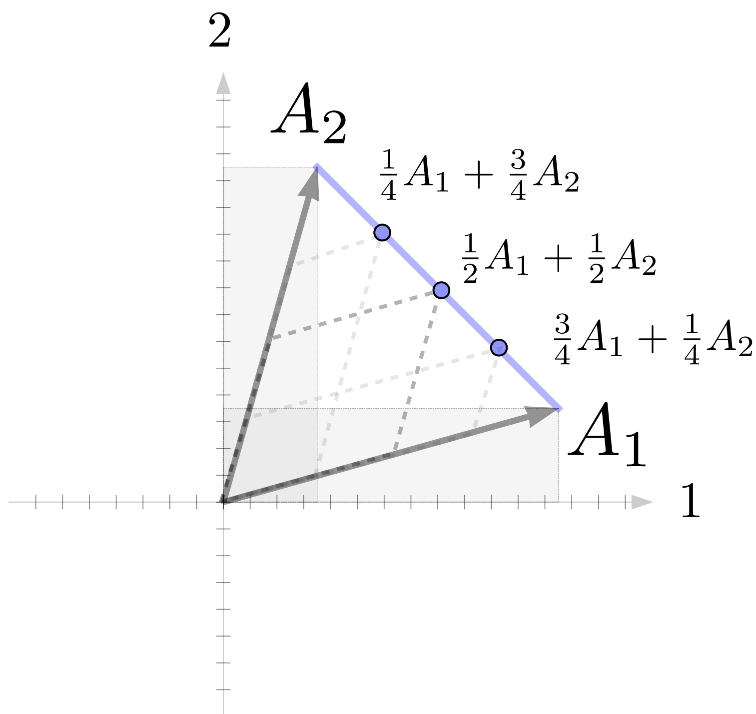


3.4 Convex Hulls

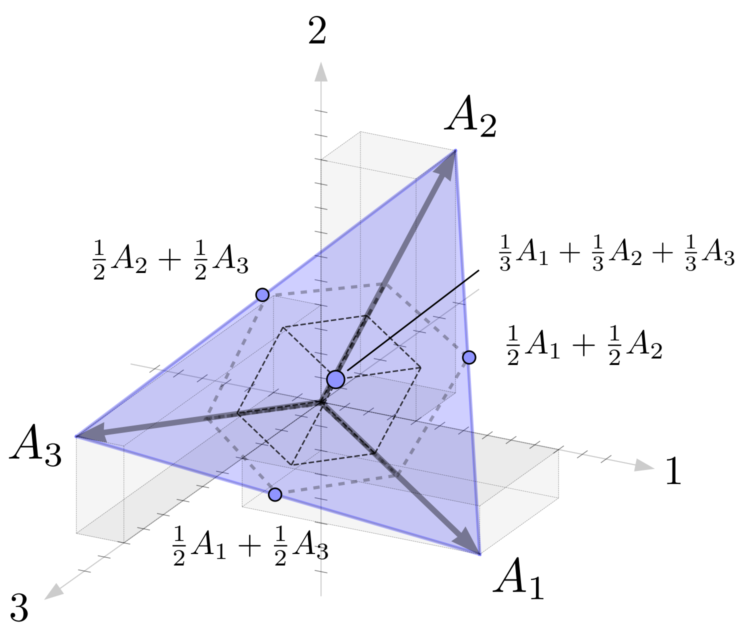
We now focus specifically on the image of the simplex, ie. the convex hull of the columns of A. We will denote this set $A\Delta$ as is fitting

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

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 $A \in \mathbb{R}^{2\times 2}$

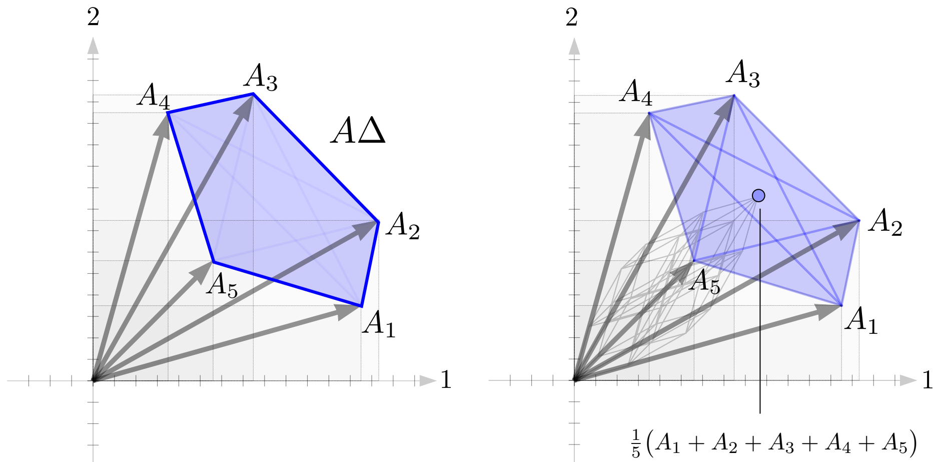


and $A \in \mathbb{R}^{3\times 3}$



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.



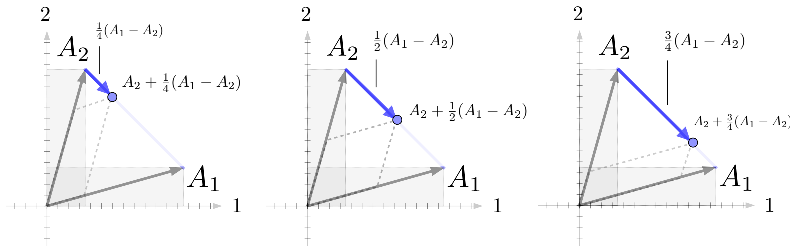
For matrices $A^{2\times 2}$ there is a specific parametrizetion of the convex hull that is often useful. In this case the set is a line segment between the two vectors and can be parametrized by a single variable α. The constraints on $x \in \mathbb{R}^2$ can be rewritten

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

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)

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

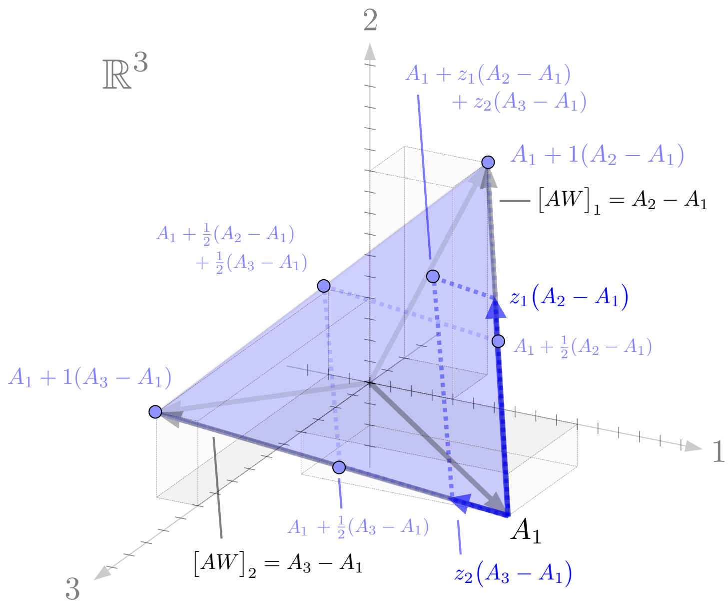
Note here that this matrix is given by AW where $W \in \mathbb{R}^{n \times n-1}$ is given by

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

The convex hull of A can be denoted

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

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



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

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

Remark 3. *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.*

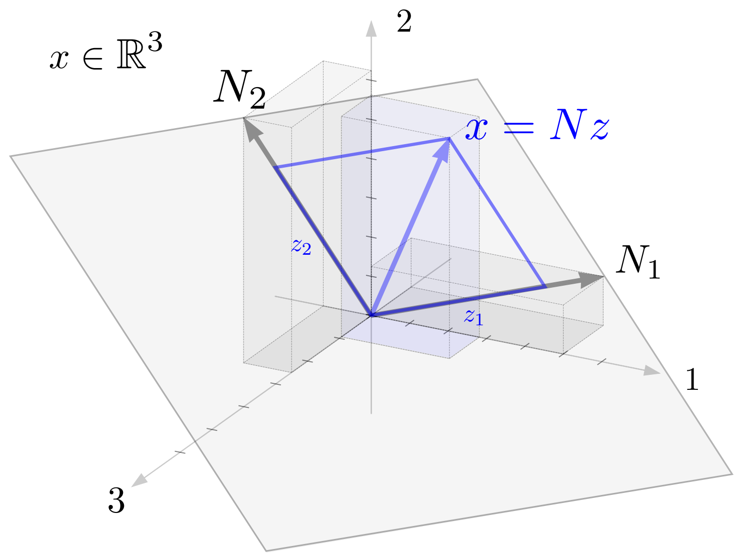
3.5 Subspace Representations

We now turn to discussing images of subspaces.

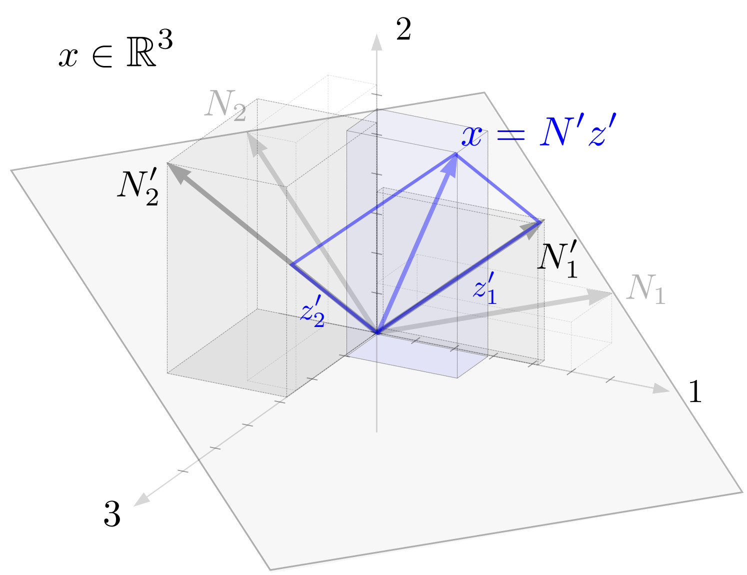
The first subspace representation

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

is naturally thought of as the range of N.



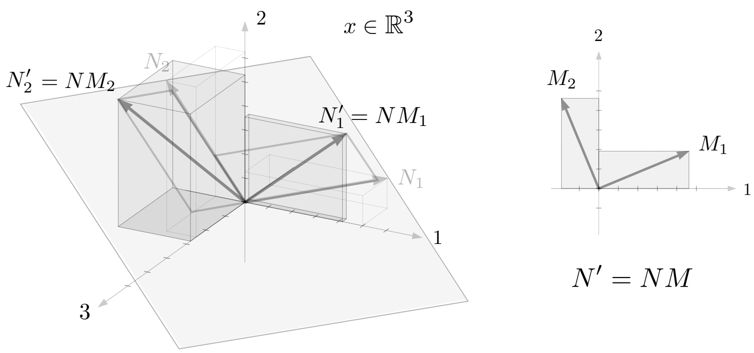
Note that for a particular subspace, the choice of N is not unique; it is only important that they have the same span. For a particular 2D subspace in $\mathbb{R}^3$, we illustrate two matrices N and N' that have the same span be used to represent the same subspace.



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

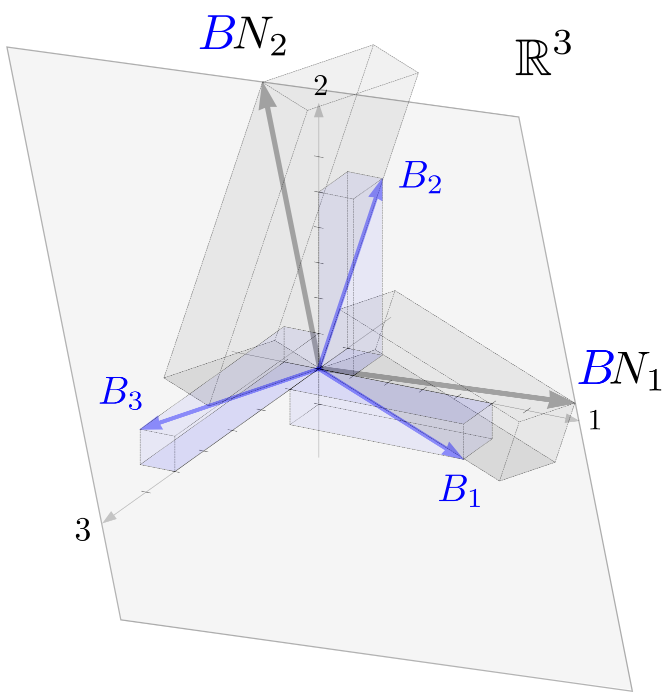
$$x = N'z' = NMz' = Nz$$

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



3.6 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.

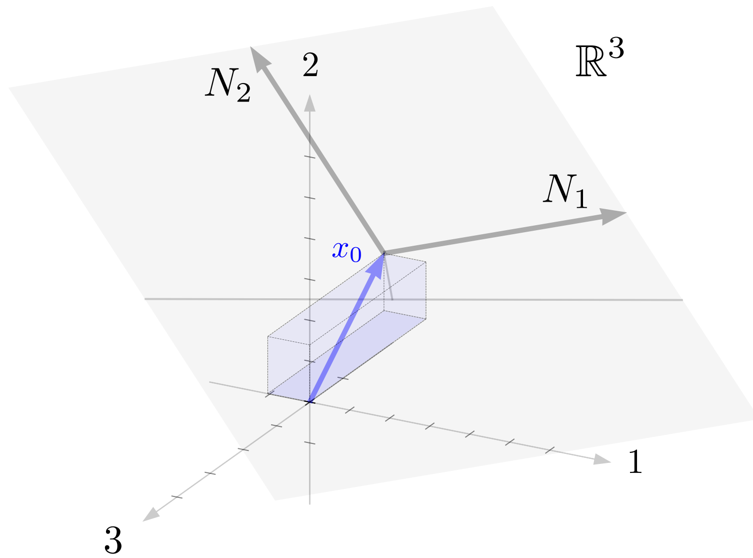


3.7 Affine Spaces

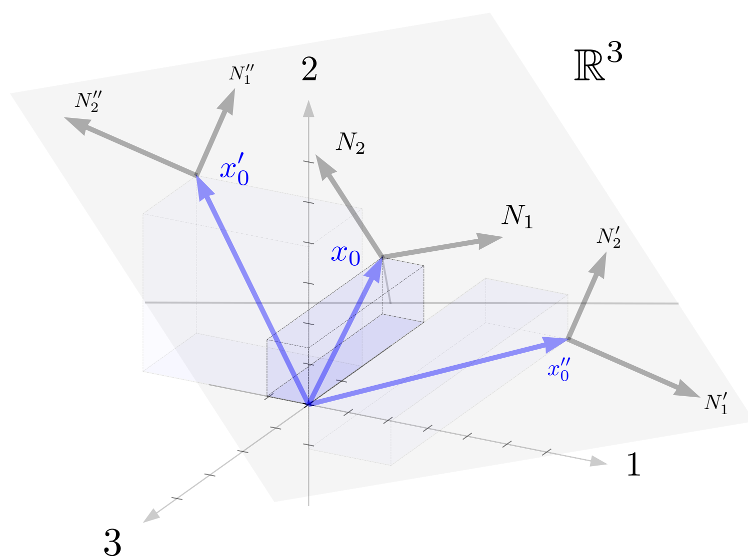
The first characterization of affine spaces

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

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

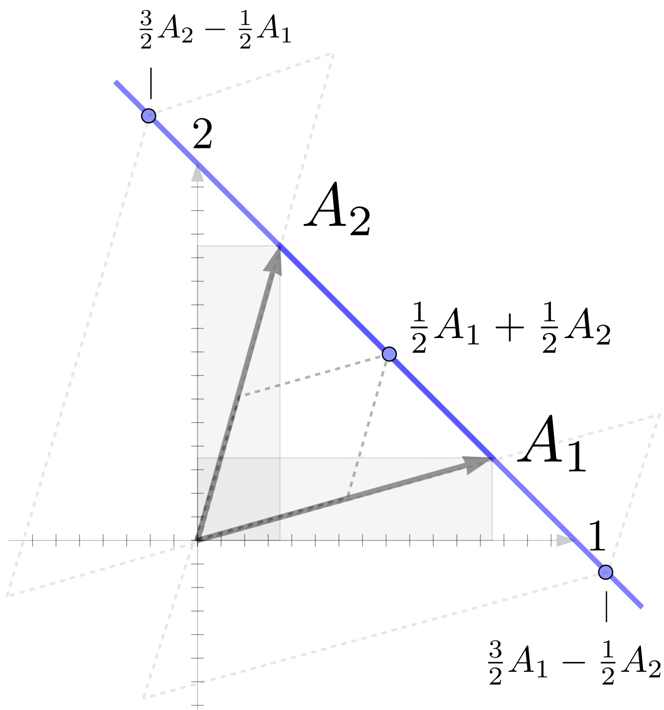
$$\begin{aligned} x &= N'z' + x_0' \\ &= NMz' + x_0' \\ &= NMz' + x_0' - x_0 + x_0 \\ &= NMz' + Nu + x_0 \\ &= N(Mz' + u) + x_0 \end{aligned}$$

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

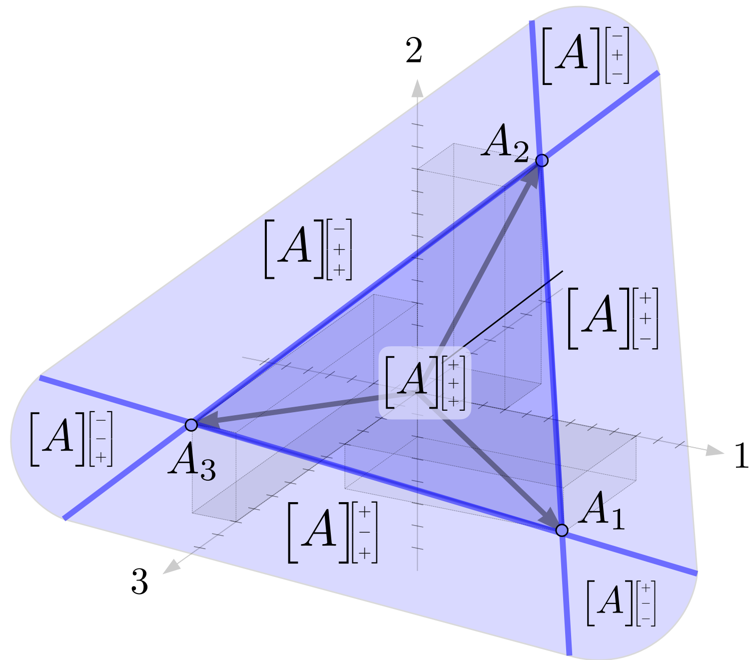
The third affine space representation

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

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

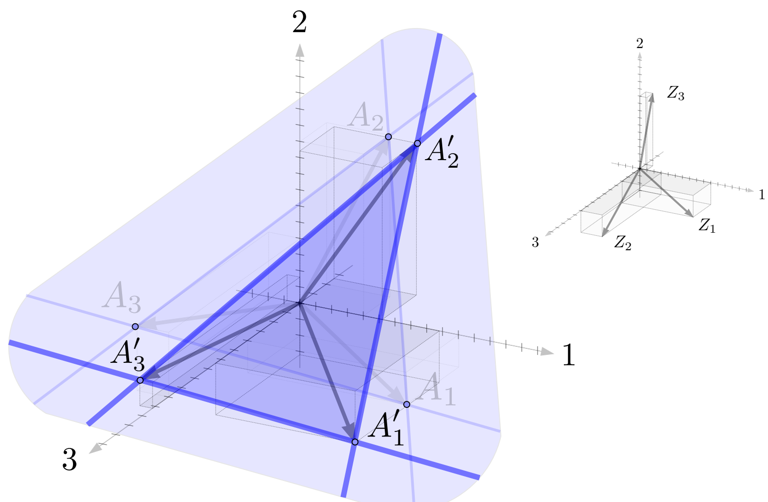


and for $A \in \mathbb{R}^{3\times 3}$ with the regions defined by z's with various sign profiles (though all summing to 1)



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

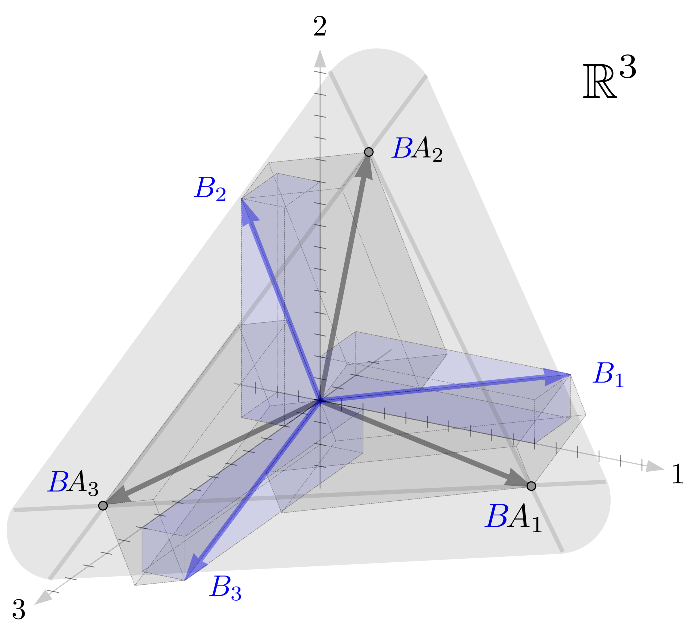
$$x = A'z' = AZz'$$

with $z = Zz'$. Note that $\mathbf{1}^T z = \mathbf{1}^T Z z' = \mathbf{1}^T z' = 1$. 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.

3.8 Affine Space Images

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.



4 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 4. *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.*

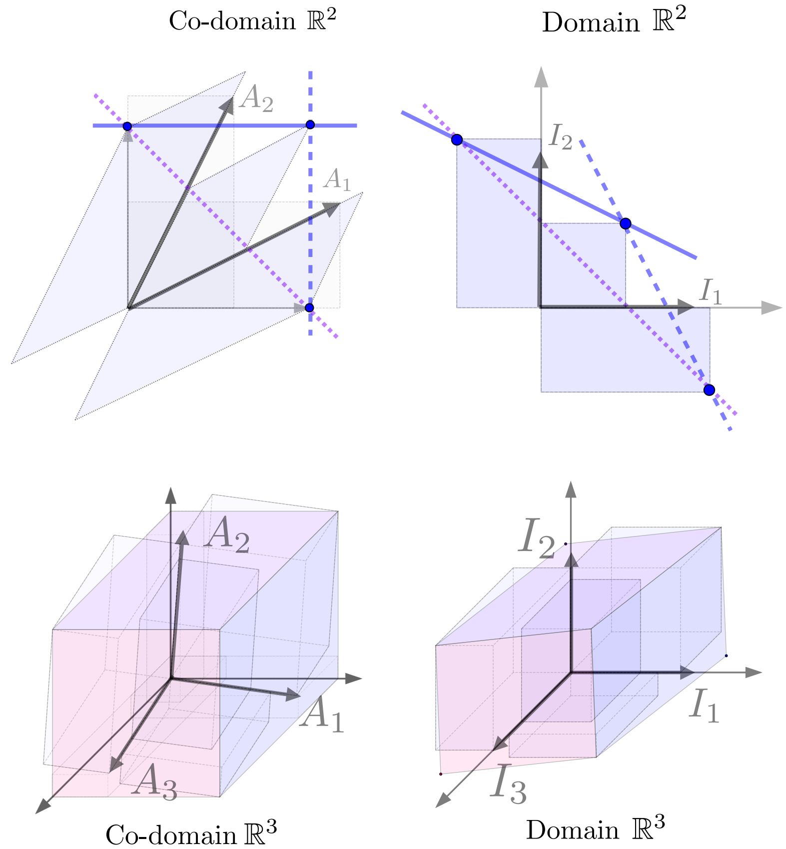
4.1 Basic Pre-Images

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.

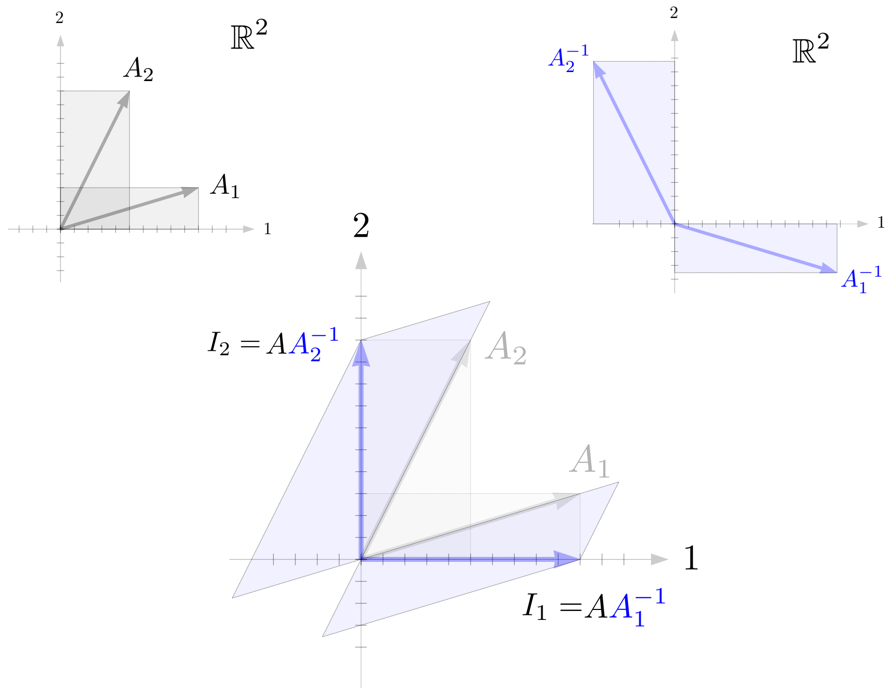
4.2 Pre-Image of Co-Domain Subspaces

While perhaps slightly more complicated, it is also very useful to visualize the pre-image of co-domain subspaces in the domain. We note that this will be especially useful for eventually discussing polytopes.

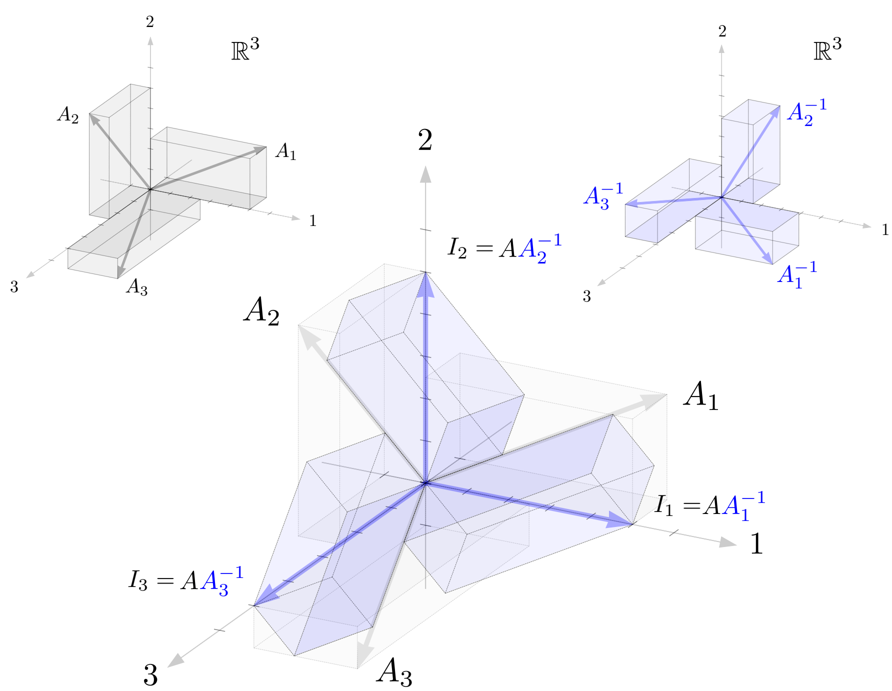


4.3 Inverses

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 $\mathbb{R}^{2\times 2}$

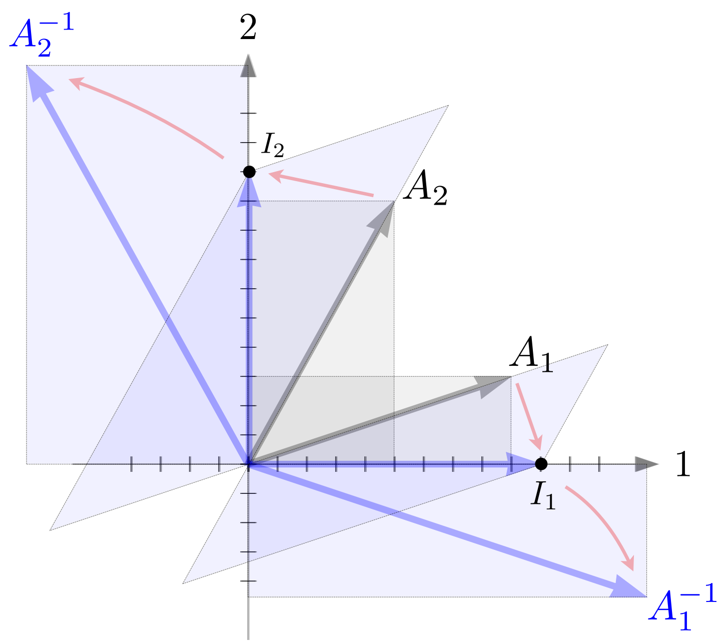


and $\mathbb{R}^{3\times 3}$



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

Remark 5. *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 2×2 example shown here.*



4.4 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

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

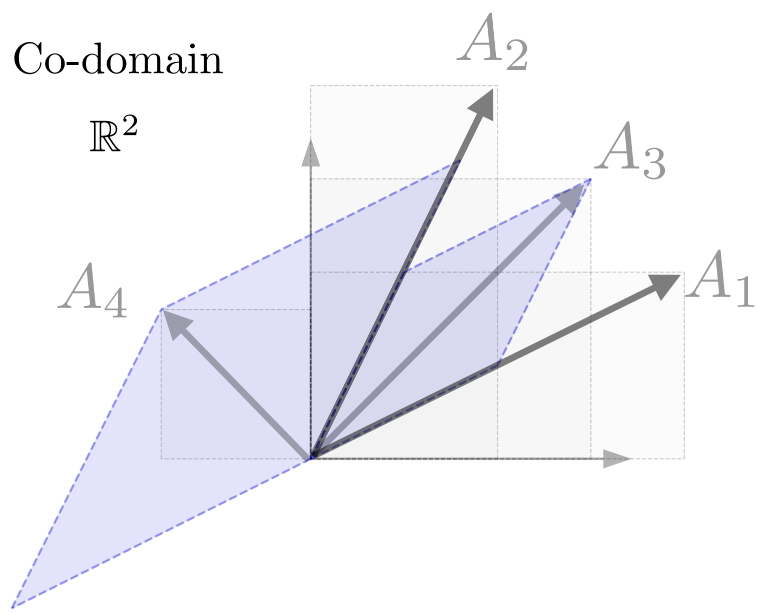
then $AN = 0$. With a little more work, we can show that the columns of N form a basis for the nullspace. The identity block proves the linear independence of the columns. We can also show explicitly that any element in the nullspace of A is in the span of N. This fact relies on the linear independence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

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

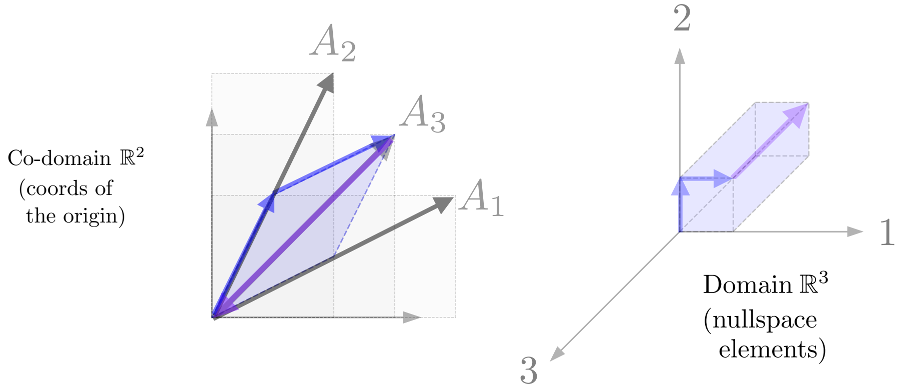
$$\text{SPAN:} \qquad Ax = 0 \Rightarrow A'\begin{bmatrix} I & B \end{bmatrix} \begin{bmatrix} x' \\ x'' \end{bmatrix} = 0$$

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

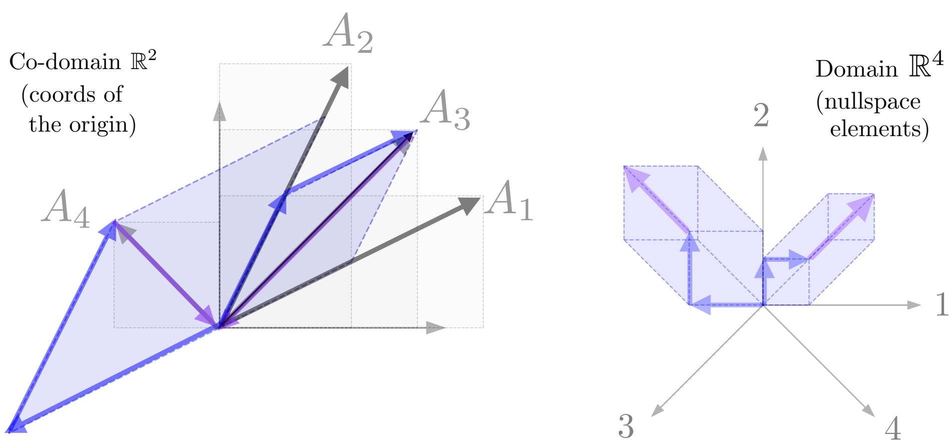
where the beginning the second line depends on the columns of A' being linearly independent. Writing the columns of A'' in terms of the columns of A' is illustrated in the figure XXX below.



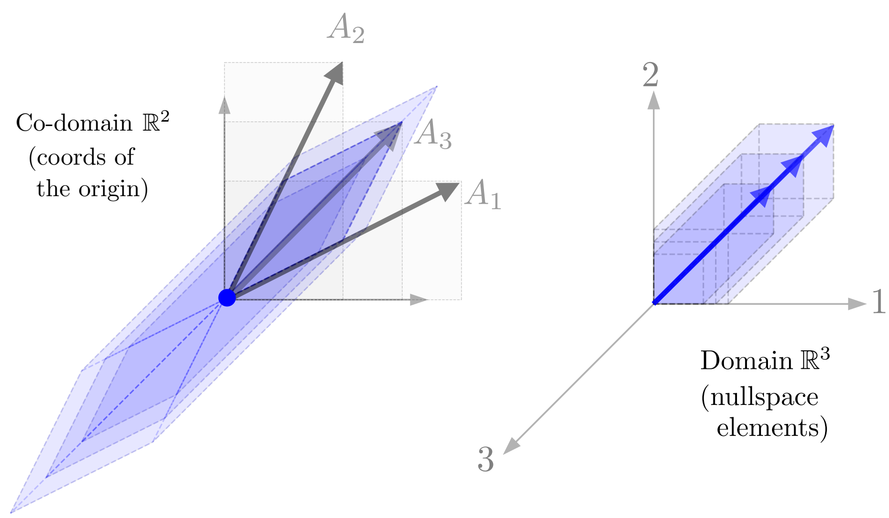
We note in this case, we are actually visualizing specific (non-zero) points in the domain that map to 0. If we expand these out in the domain we get the points in these figure illustrated for a matrix with one "extra" column $A \in \mathbb{R}^{2\times 3}$



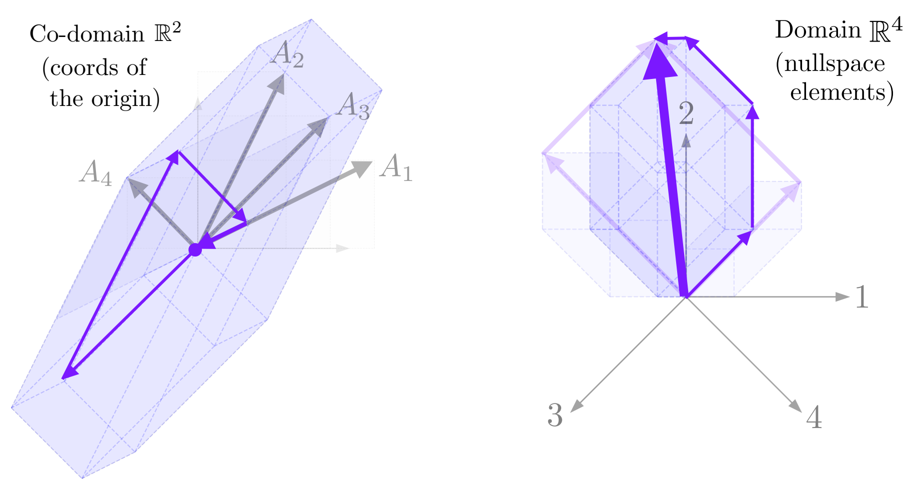
and for a matrix with two extra columns $A \in \mathbb{R}^{2\times 4}$.



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.

4.5 Range of A^T

The last and most difficult subspace that we will consider in reference to column geometry is the span of the rows of the matrix, the range of A^T. It is perhaps a surprising fact of matrix geometry that while the algebraic relationship between columns and rows is quite simple, any geometric relationships are much more subtle. We offer several perspectives here on how to "see" the rows via the column geometry of a matrix. The authors would like to note this section is experimental and they find it the least satisfactory in this paper.

4.5.1 Input Directions

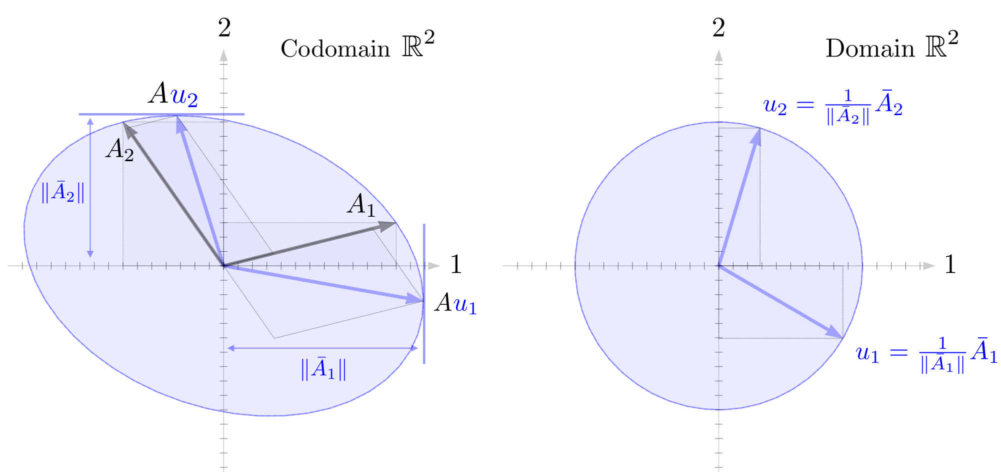
Another option for row visualization is to note that each row defines an input direction that provides the maximum gain in a particular coordinate of the co-domain. To be more specific, consider a unit vector u in the domain, $u \in \mathbb{R}^n$ and multiply u by a matrix A, Au.

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

The ith coordinate of the output is given by $\bar{A}_i^T u$, the inner product between the ith row and u. The quantity $\bar{A}_i^T u$ is maximized when $u = \frac{1}{\|\bar{A}_i\|}\bar{A}_i$, ie. when u points in the same direction as row $\bar{A}_i^T$. Note also that when this is the case,

$$\bar{A}_i^T u = \frac{\bar{A}_i^T \bar{A}_i}{\|\bar{A}_i\|} = \|\bar{A}_i\|,$$

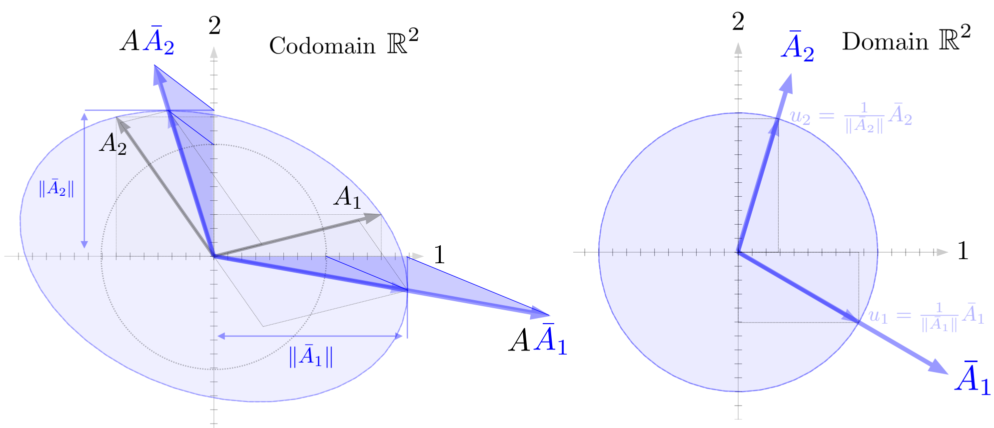
ie. $\bar{A}_i^T u$ is the magnitude of row $\bar{A}_i^T$. This leads to the following visualization. Visualize the image of the unit sphere (in the domain) via the column geometry. The direction (in the domain) that gives the maximum output along each axis is the direction of that row. The gain along that axis is the magnitude of the row. This illustrated for $A \in \mathbb{R}^{2\times 2}$.



This illustration separates the action of the row into it's magnitude and direction. The direction is shown in the domain, while the magnitude is shown in the codomain. Algebraically, this is equivalent to separating A in the following way

$$A = \begin{bmatrix} - & \bar{A}_1^T & - \\ & \vdots & \\ - & \bar{A}_m^T & - \end{bmatrix} = \underbrace{\begin{bmatrix} \|\bar{A}_1\| & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \|\bar{A}_m\| \end{bmatrix}}_{\text{dg}} \begin{bmatrix} - & \frac{1}{\|\bar{A}_1\|}\bar{A}_1^T & - \\ & \vdots & \\ - & \frac{1}{\|\bar{A}_m\|}\bar{A}_m^T & - \end{bmatrix}$$

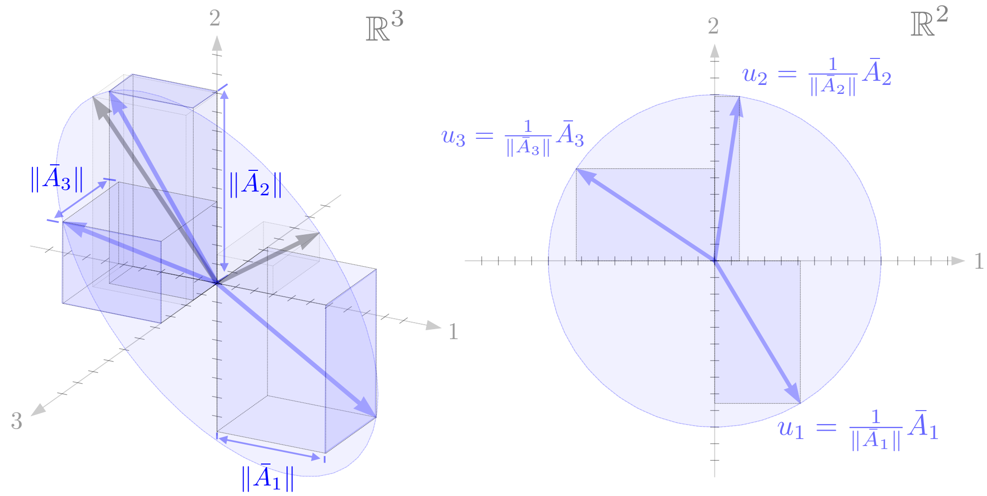
We can illustrate the magnitude in the domain as well; however the method may be less than satisfying. We do this by applying the stretching that moves the unit circle (in the domain) to the matrix dg to each of the rows in the domain. This can be illustrated using similar triangles as follows.



Again, we note that this illustration is slightly unsatisfying in that the stretching is difficult to visualize. For the remainder of the examples in this section, we will simply illustrate the unit vectors in the domain (the directions of the rows) and magnitudes in the co-domain as in the first example above. Perhaps an interested reader can provide a more elegant illustration. We now consider cases where A is not square.

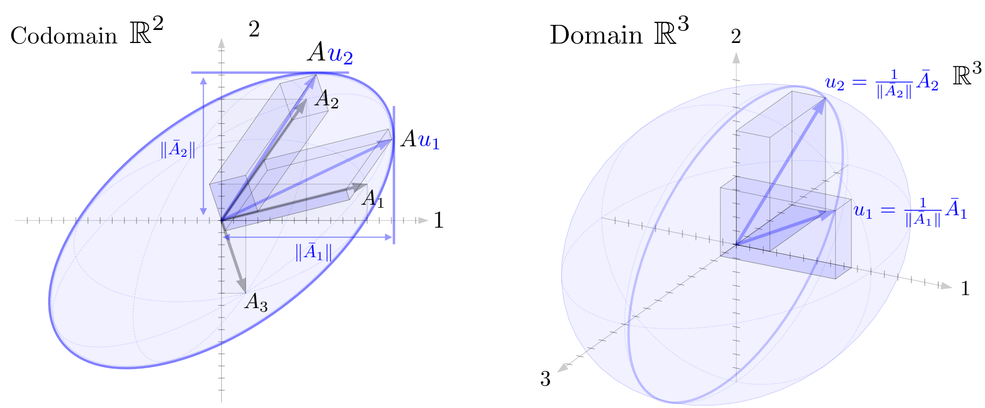
4.5.2 Input Directions - Tall Matrices

When A is a tall matrix the same illustration applies however the image of the domain unit circle has lower dimension than the codomain. We illustrate this here for $A \in \mathbb{R}^{3\times 2}$. Note that there are three points on the 2D unit circle that maximize the distance along each separate axis in the codomain $\mathbb{R}^3$ that correspond to each row of the matrix.

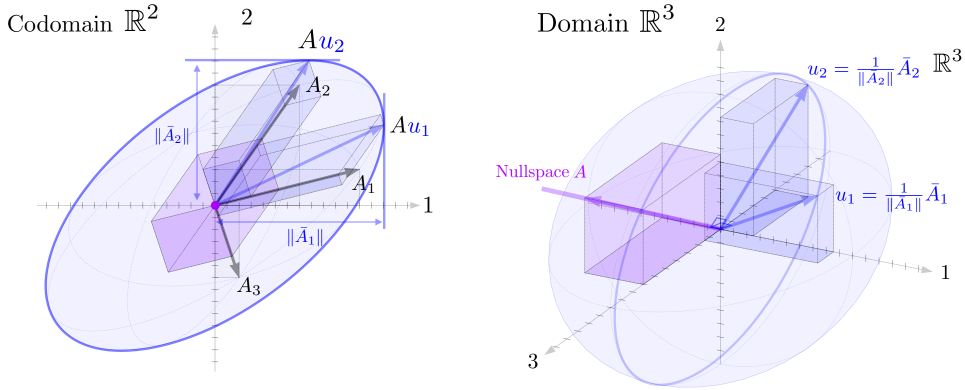


4.5.3 Input Directions - Fat Matrices

For fat matrices, the domain unit circle has larger dimension and the action of the matrix collapses one or more of those dimensions (the nullspace of A). Here the rows are the points on the unit circle that maximize the gain in a particular direction and also do not contain a component in the nullspace. We also illustrate this here for $A \in \mathbb{R}^{2\times 3}$.

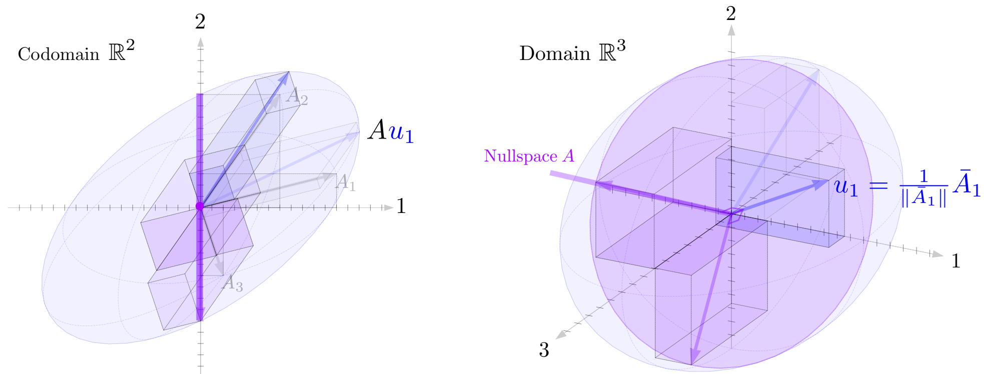


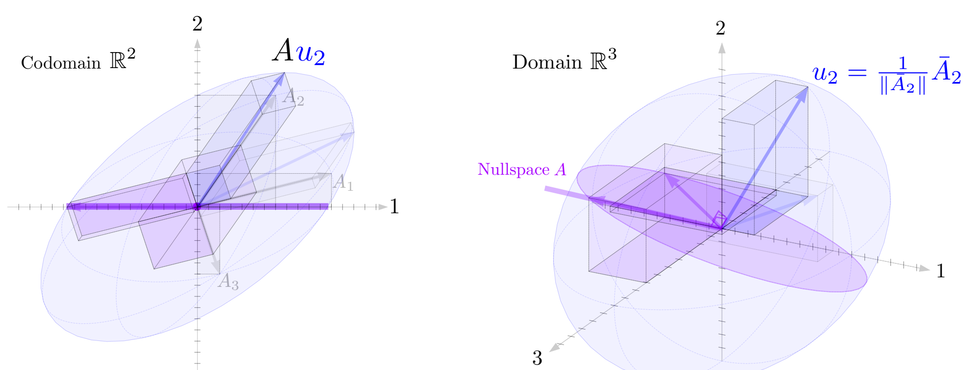
We can also visualize the nullspace in a similar manner to be the direction(s) on the unit sphere that are do not show up in the image. Of course, these are orthogonal to the rows.



This visualization idea could also be relaxed to visualize how close u lines up to each column. If the output $\bar{A}_i^T u$ is far in the positive direction along axis

Another quite fruitful modification of this idea is to consider the subspaces orthogonal to each row. The subspace orthogonal to the ith row of A is the pre-image of the set with 0 along the ith axis (in the co-domain). This pre-image is illustrated here relative to the unit sphere in the domain as well as the subspace containing it.





4.6 Affine Spaces (Domain)

The nullspace discussion above extends to visualizing sets of the form

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

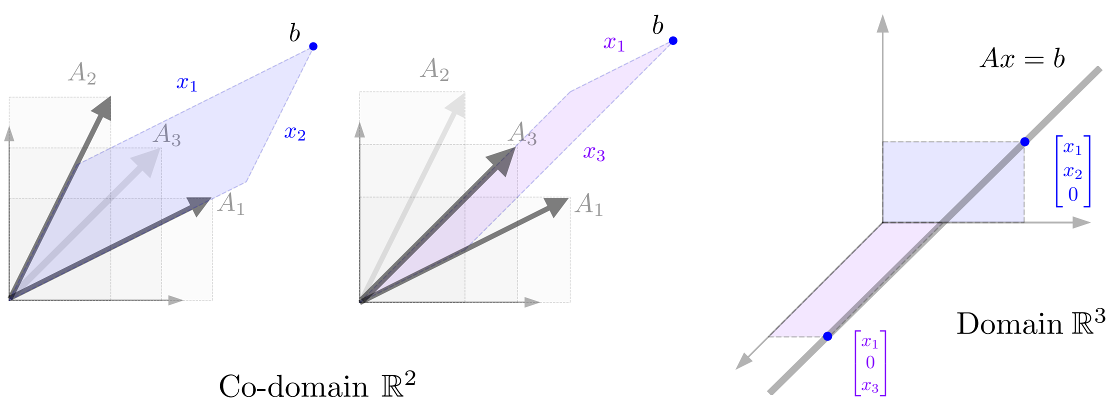
for $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and $k < n$. These type of sets can be more explicitly characterized using a basis for the nullspace of A written as the columns of the matrix $N \in \mathbb{R}^{n\times k}$ using the form

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

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.



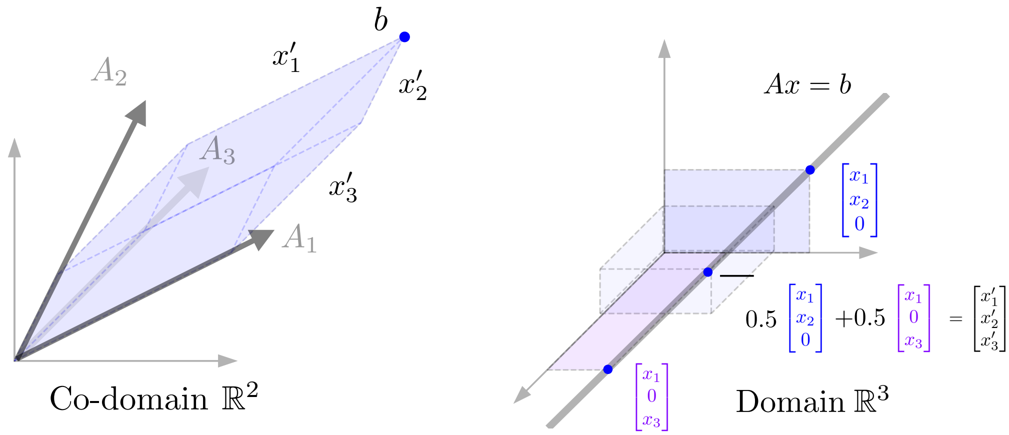
The line between these two points can be defined as the set of points

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

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



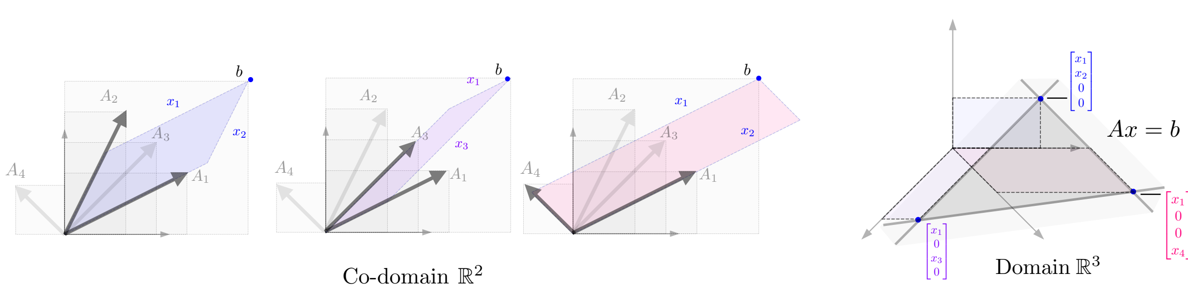
Further more we can see that the difference between these two points (or any scalar multiple of this difference) is in the nullspace of A. Again explicitly

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

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

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



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

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 $Ax = b$.

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

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.

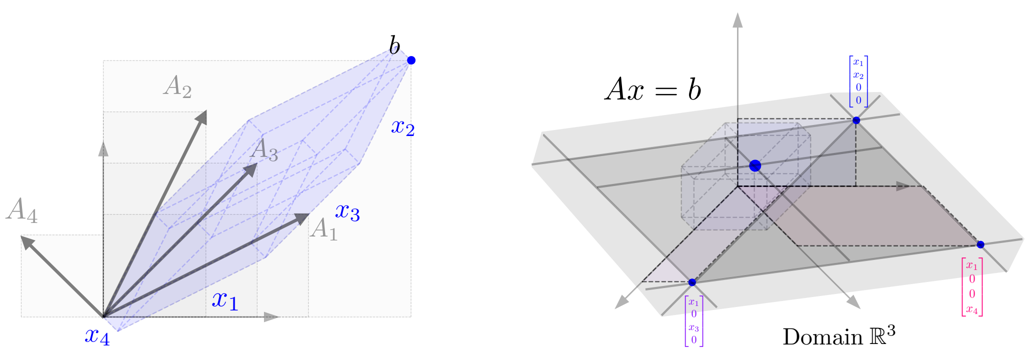
In general, the affine space given can be characterized by a specific solution x_0 plus any element in the nullspace. The solution x_0 can be chosen to be any of the elements given above. Often it is chosen to be the minimum norm solution, This is computed as

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

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)