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

Dan Calderone

***Abstract*—This tutorial paper gives basic visualizations for real matrices in $\mathbb{R}^{2\times 2}$ wih a primary focus on column geometry. Basic notation and basic column and row geometry are given followed by visualizations of several basic types of matrices. Matrix multiplication is discussed. Transposes and inverses are discussed with a focus on symmetric/skew-symmetric matrices and visualization. Similarity transforms and eigenvalue decompositions are discussed. Explicit algebraic characterization of eigenvalues/eigenvectors along with thorough visualizations. Special attention is then given to complex eigenvalues/eigenvectors including a discussion of rotation matrices. Symmetric and definite matrices are discussed in the context of quadratic forms. Grammians (shape matrices) are discussed; polar decompositions are derived and discussed, and the singular value decomposition is discussed in detail with analogies drawn from complex numbers. The final section of the paper details linear vector fields. The spectral mapping theorem and the matrix exponential are discussed along with stability criteria in continuous and discrete time.**

I. Introduction

Real matrices in $\mathbb{R}^{2\times 2}$ show up in every corner of modern mathematics. Beyond being useful in their own right, they also provide foundational examples and intuition for studying invertible linear maps (square matrices) in general. In this paper, we give many detailed visualizations of basic structural results about $\mathbb{R}^{2\times 2}$.

The initial section covers basic notation as well as gives basic details of column and row geometry and the image of sets under 2×2 linear transformations. Our analysis in this paper will focus primarily on column geometry as it is the most natural but row geometry will be discussed as well at several points. Matrix multiplication is discussed briefly as well. We next detail the structure of several basic classes of square 2×2 matrix structures including diagonal, upper/lower triangular, symmetric/skew-symmetric, rotations/reflections, and nilpotent structures. This section is meant to give a general flavor and build basic spatial intuition.

The next section discusses the geometry of the matrix transpose, ie. the geometry of the rows relative to the columns. While algebraically immediate, this geometry is actually fairly subtle. Particular attention is given to the symmetric and skew-symmetric portions of the matrix. Inverses are then discussed. Whereas transpose are algebraically simple and geometrically complicated, inverses have the opposite flavor (geometrically simple but algebraically complicated).

We then turn our attention to the rich subject of similarity transformations and eigenvalue decompositions. Similarity transforms are visualized with specific attention given to orthonormal similarity transforms (similarity transforms that are also congruent). The characteristic polynomial and it's relation to eigenvalues and left/right eigenvectors is thoroughly visualized and discussed. Formulas for eigenvalues, eigenvectors, and diagonalizations are given. Special attention is given to the complex eigenvalue case and pseudo-diagonal/rotational forms of complex eigenvalue decompositions. We also include a discussion of repeated eigenvalues, Jordan form, and nilpotent matrices. We then present in detail how eigenvalues and both left/right eigenvectors relate to the column geometry of a matrix in the both the real and complex eigenvalues cases. These particular visualizations are detailed and extensive. The complex case is then expanded further to detail it's rotational structure and specific attention is given to true rotation/reflection matrices. We also give specific attention to skew-symmetric matrices as real matrices with purely imaginary eigenvalues. Finally, we conclude the initial eigen-decomposition discussion with a brief discussion of the spectral mapping theorem. We next turn our attention to symmetric matrices in the context of quadratic forms. Quadratic form surfaces and their relationships to symmetric matrix eigenstructure is discussed. Positive-definite, negative-definite, and indefinite matrices are discussed.

The next section of the paper contains a detailed discussion of matrix shapes including the polar decomposition and singular value decomposition. The two Gramian matrices and, more importantly, their square roots, are discussed as the primary two definitions of matrix's positive definite "shape". From there we derive and visualize the polar decomposition in both contexts. Finally, we use the eigen-structure of the Gramian matrices to give the singular value decomposition (the classical construction) and give visualizations. Detailed connections between each of these decompositions as well as the sym/skew-sym decomposition are given as a thorough discussion of analogies with complex numbers in their Cartesian and polar form.

The final section of this paper details the structure of linear vector fields (ordinary differential equations) in the linear time invariant case. Basic solutions in the form of the matrix exponential in continuous and discrete time are given. The relationship between eigen-structure and trajectories is detailed. Stability criteria in both continous and discrete time are given in terms of eigenvalues as well as various parametric tests for stability. Some of these are classical results while others are somewhat novel.

Remark 1. *The primary section missing from this paper is perhaps one focusing on matrix commutators. The authors hope to add this section at some point in the future.*

A. *Prerequisites and Follow-ups*

The authors have created an interactive tool for exploring the geometries in this paper presented here:

https://danjcalderone.github.io/dcmath/linalg/2x2visualizer.html

It can be read on it's own without much difficulty; however it does assume a familiarity with the notation and vector/matrix visualization techniques presented in the following monographs.

- Vector visualizations:
  https://danjcalderone.github.io/papers/vectors.pdf
- Column geometry:
  https://danjcalderone.github.io/papers/columns.pdf

This paper is also meant to be part one of a three part series. The second paper expands many of these results/visualizations to real matrices in $\mathbb{R}^{3\times3}$; the third paper discusses extensions to general matrices in $\mathbb{R}^{n\times n}$ with visualizations given in $\mathbb{R}^{4\times4}$. This last paper is (of course) far less thorough since the space $\mathbb{R}^{n\times n}$ is a vast mathematical landscape that has never been fully explored. Any "thorough" discussion would have to include countless specific types of matrices. The visualizations in this last paper are also only meant to be experimental and to give a flavor for how an ambitious student of visualization might seek to extend the the visualization techniques in the first two papers to higher dimensional geometries. As such, they should only be viewed in parallel to the first two monographs. The authors also take no responsibility for any confusion that may result from viewing them. The dissatisfied reader is always heartily encouraged to make improvements or re-fall in love with pure algebraic insight. [1]

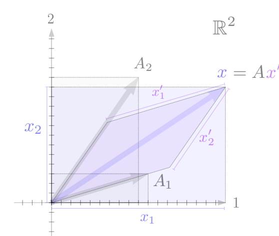
II. Basic Column Geometry

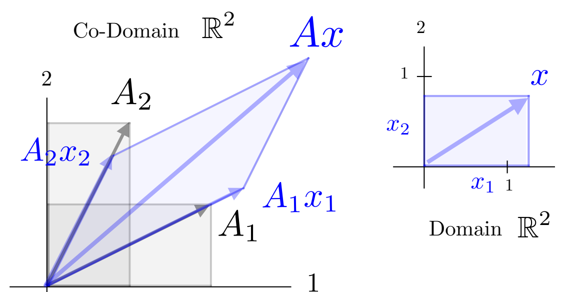
The columns of a matrix $A \in \mathbb{R}^{m\times n}$ are vectors in the co-domain of the linear map. For a matrix $A \in \mathbb{R}^{2\times2}$, we can split the matrix into two columns or two rows depending on context.

$$A = \begin{bmatrix} | & | \\ A_1 & A_2 \\ | & | \end{bmatrix} = \begin{bmatrix} - & \bar{A}_1^T & - \\ - & \bar{A}_2^T & - \end{bmatrix}$$

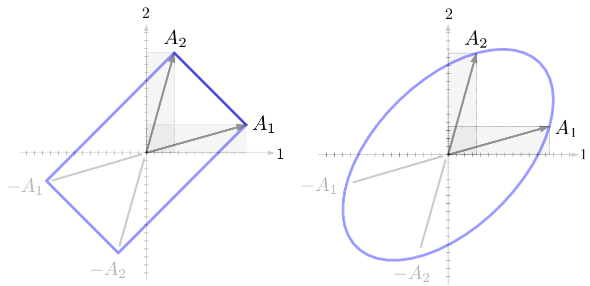
Each individual column $A_j \in \mathbb{R}^2$ tells where the jth standard basis vector (in the domain) gets mapped under the transformation. Explicitly $AI_j = A_j$. We can see where a vector $x \in \mathbb{R}^n$ in the domain gets mapped by breaking up x into a linear combination of standard basis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

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



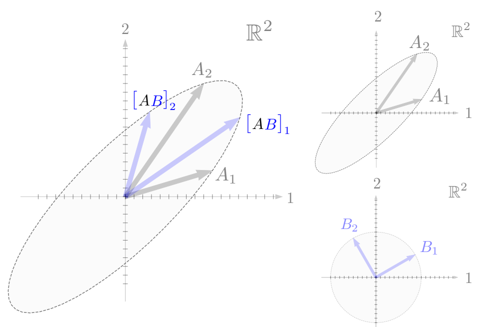


This column geometry picture can be applied to sets of vectors at a time to see their image through the map in the codomain. This can be quite useful in getting a sense for how a linear map distorts the space. Images:

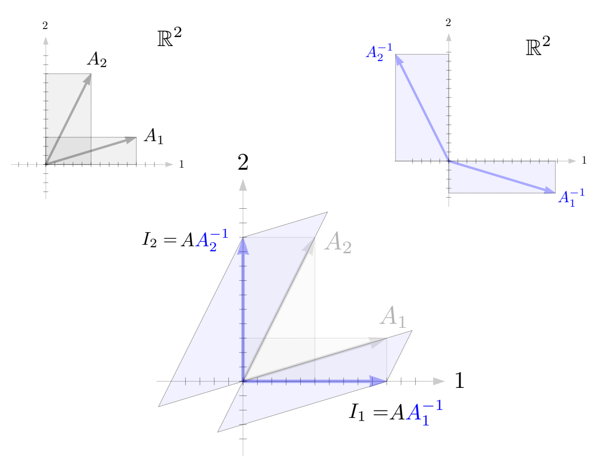


III. Matrix Multiplication & Inverses

It can also be used to visualize composition of matrix matrices. For matrix product AB, drawing the columns of B relative to A (as the new axes) gives the columns of AB in the codomain of the composition map. We give an illustration here where B is rotation matrix to illustrate the feel for this geometry.



This technique is also surprisingly useful for illustrating the columns of a matrix inverse. A matrix inverse is the matrix B such that AB maps to the identity. For any 2×2 matrix this can be read off fairly quickly simply by drawing the columns of A as illustrated in this image.



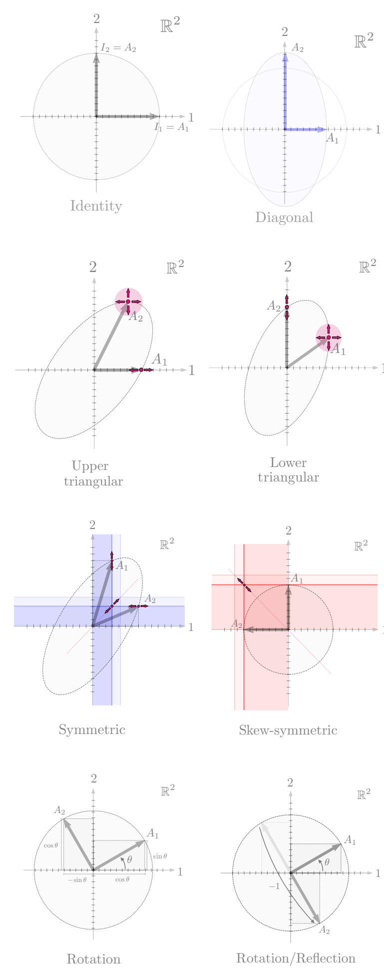
For the sake of completeness, we also include the algebraic formula for the inverse here.

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

A few more comments about inverses are detailed in the section.

IV. Matrix Types

To get a further feel for matrix column geometry, we briefly illustrate the columns of several specific types of matrices. This list is not meant to be exhaustive and the reader is strongly encouraged to explore other matrix types of interest via the column geometry picture. (This idea is even richer for matrices larger than 2×2 with many more types of matrices creating interesting geometries).



V. $MHPK$ Decomposition

The matrix A can be decomposed as

$$A = \begin{bmatrix} m+h & p-k \\ p+k & m-h \end{bmatrix}$$

where

$$m = \tfrac{1}{2}(a+d), \quad h = \tfrac{1}{2}(a-d),$$
$$p = \tfrac{1}{2}(b+c), \quad k = h = \tfrac{1}{2}(c-b)$$

Note that here m and p are the averages of the diagonal and off-diagonal elements (respectively) and h and k are the differences. This $mhpk$-parametrization is particularly useful for considering limiting cases or special types of matrices. We detail some of these in the image below. Only $m \neq 0$ is a scaled identity matrix; $m, h, p \neq 0$ is symmetric. only $k \neq 0$ is skew symmetric; only $m, k \neq 0$ is a scaled rotation; $h, p \neq 0$ is symmetric with zero trace. Other interesting categories of 2×2 matrices such as $m, p \neq 0$ (symmetric matrices with constant diagonal) or $h, k \neq 0$ (a zero-trace matrix) that has some rotation like properties could be considered as well.

We comment also that this parametrization suggests several matrix decompositions. Specifically, we will consider

$$A = \underbrace{\begin{bmatrix} m & -k \\ k & m \end{bmatrix}}_{M} + \underbrace{\begin{bmatrix} h & p \\ p & -h \end{bmatrix}}_{H}$$

Often it will be useful to enumerate the columns of each as well.

$$M = \begin{bmatrix} m & -k \\ k & m \end{bmatrix} = \begin{bmatrix} | & | \\ u_+ & u_- \\ | & | \end{bmatrix}$$
$$H = \begin{bmatrix} h & p \\ p & -h \end{bmatrix} = \begin{bmatrix} | & | \\ v_+ & v_- \\ | & | \end{bmatrix}$$

We compute the norms of u and v as they will show up repeatedly in what follows.

$$|u| := \|u_+\| = \|u_-\| = \sqrt{m^2 + k^2}$$
$$|v| := \|v_+\| = \|v_-\| = \sqrt{p^2 + h^2}$$

Here we've decomposed A into two orthogonal matrices, M and H. The first M is a scaled rotation and the second H is a symmetric zero-trace matrix that turns out to be a scaled reflection. We list two propositions that make the above statements explicit. The proofs (omitted) are simply direct computation (with application of trignometric identities).

Proposition 1 (M: Scaled Rotation)**.**

$$\begin{bmatrix} m & -k \\ k & m \end{bmatrix} = |u| \begin{bmatrix} \frac{m}{|u|} & -\frac{k}{|u|} \\ \frac{k}{|u|} & \frac{m}{|u|} \end{bmatrix} = |u| \underbrace{\begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}}_{R_\phi}$$

with $\phi = \arctan(k/m)$

Proposition 2 (H: Diagonalization)**.**

$$\begin{bmatrix} h & p \\ p & -h \end{bmatrix} = \underbrace{\begin{bmatrix} \cos\frac{\psi}{2} & -\sin\frac{\psi}{2} \\ \sin\frac{\psi}{2} & \cos\frac{\psi}{2} \end{bmatrix}}_{R_{\psi/2}} \begin{bmatrix} |v| & 0 \\ 0 & -|v| \end{bmatrix} \underbrace{\begin{bmatrix} \cos\frac{\psi}{2} & \sin\frac{\psi}{2} \\ -\sin\frac{\psi}{2} & \cos\frac{\psi}{2} \end{bmatrix}}_{R^T_{\psi/2}}$$

with

$$R_{\psi/2} = \begin{bmatrix} \cos\frac{\psi}{2} & -\sin\frac{\psi}{2} \\ \sin\frac{\psi}{2} & \cos\frac{\psi}{2} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \sqrt{|v|+h} & -\sqrt{|v|-h} \\ \sqrt{|v|-h} & \sqrt{|v|+h} \end{bmatrix}$$

and $\psi = \arctan(p/h)$

More propositions about M and H are given in the eigenvalue decomposition section, the singular value decomposition section and also the appendices.

This decomposition will prove surprisingly useful in illuminating the geometry of both the EVD and SVD of a 2×2. Much of the inspiration for this direction comes from noting the following formulas for the trace, determinant, and discriminant (Δ) (from the characteristic polynomial) of the matrix.

$$\begin{aligned} \frac{1}{2}\text{tr}(A) &= m \\ \det(A) &= ad - bc \\ &= (m+h)(m-h) - (p+k)(p-k) \\ &= m^2 + k^2 - h^2 - p^2 \\ \sqrt{\Delta} &= \sqrt{\left(\frac{\text{tr}(A)}{2}\right)^2 - \det(A)} \\ &= \sqrt{m^2 - (m^2 + k^2 - h^2 - p^2)} \\ &= \sqrt{h^2 + p^2 - k^2} \end{aligned}$$

We can write the determinant and discriminant formulas in terms of the columns of M and H as well.

$$\det(A) = |u|^2 - |v|^2$$
$$\sqrt{\Delta} = \sqrt{|v|^2 - k^2}$$

Finally, we return the matrix inverse formula and write it in the $mhpk$ coordinates.

$$A^{-1} = \frac{1}{det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{det(A)} \begin{bmatrix} m-h & -p+k \\ -p-k & m+h \end{bmatrix}$$

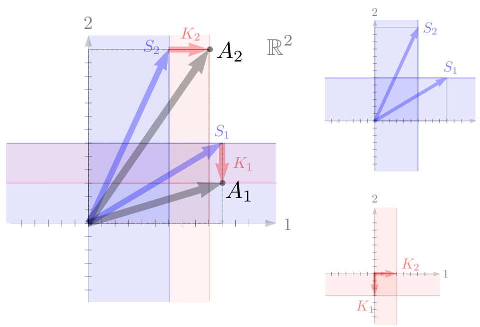
It is interesting to note that (along with scaling by the determinant) the matrix inverse operation negates h, p, and k, but leaves m untouched.

VI. Sym-Skew Decomposition & Transposes

We note that the above $M - H$ decomposition is different from the more traditional symmetric/skew-symmetric decomposition of a matrix. Expressed here in the $mhpk$ coordinates, the symmetric/skew-symmetric decomposition is given by

$$
\begin{aligned}
A &= \underbrace{\tfrac{1}{2}(A + A^T)}_{S} + \underbrace{\tfrac{1}{2}(A - A^T)}_{K} \\
&= \underbrace{\begin{bmatrix} m+h & p \\ p & m-h \end{bmatrix}}_{S} + \underbrace{\begin{bmatrix} 0 & -k \\ k & 0 \end{bmatrix}}_{K}
\end{aligned}
$$

The primary difference is whether or not the component mI is combined with the k term or the h, p terms. We can illustrate the column geometry of this decomposition as follows.



The symmetric and skew-symmetric parts of a matrix act in many ways like the real and imaginary parts of a complex number and the symmetric/skew-symmetric decomposition is analogous to the cartesian representation

$$
\begin{aligned}
\text{Complex Number:} \quad & z = \underbrace{a}_{\text{Real}} + \underbrace{bi}_{\text{Imag}} \\
\text{Matrix:} \quad & A = \underbrace{S}_{\text{Sym}} + \underbrace{K}_{\text{Skew-Sym}}
\end{aligned}
$$

This analogy is rich and subtle but also not as complete as one might hope given the fact that multiplying by a matrix can be much more complicated than multiplying by a complex number. Symmetric matrices have real eigenvalues; skew-symmetric matrices have imaginary eigenvalues. Symmetric matrices represent pure stretchings; skew-symmetric matrices represent pure rotations. But in general, the eigenvalues of the full matrix are not the eigenvalues of the symmetric part plus the eigenvalues of the skew symmetric part and adding any component bi to a real number makes it complex, but a matrix with a non-zero skew-symmetric part can still have only real eigenvalues.

Mathematically, the complications arise from the fact that in general the symmetric and skew-symmetric parts don't commute. In the 2×2 case, K commutes with mI but anti-commutes with H.

$$
mIK = KmI, \qquad KH = -HK
$$

From this we get, that the complex number analogy is very strong when $H = 0$. Indeed in this case, we have simply

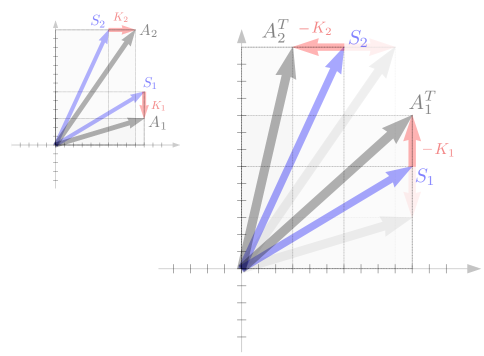
$$
A = M = \begin{bmatrix} m & -k \\ k & m \end{bmatrix}
$$

and as a scaled rotation M has eigenvalues of $m \pm ki$. mI represents a pure scaling; the addition of a non-zero H changes the shape of the stretching done by the symmetric part of A and thus breaks the tight analogy. More details on these ideas are given in the eigenvalue section and more analogies with complex numbers are explored in the sections on polar decomposition and SVD.

The $mhpk$ coordinates do provide an interesting geometric perspective on the transpose operation. The transpose can be obtained by negating the skew-symmetric part of the matrix

$$
A^T = S - K
$$

which amounts to simply negating the k term which we illustrate here.



VII. Eigenvalue Decomposition

The characteristic polynomial of the matrix is given by

$$
\begin{aligned}
\chi_A(s) &= s^2 - tr(A) + det(A) \\
&= s^2 - (a+d)s + (ad - bc)
\end{aligned}
$$

Note in the $mhpk$-parametrization this becomes

$$
\chi_A(s) = s^2 - 2ms + m^2 - h^2 - p^2 + k^2
$$

A. *Eigenvalues*

The eigenvalues are then given by the roots of the characteristic polynomial which in this case can be computed using the quadratic equation.

$$\lambda_{1,2} = \tfrac{1}{2}\mathrm{tr}(A) \pm \sqrt{\left(\tfrac{\mathrm{tr}(A)}{2}\right)^2 - \det(A)}$$

$$\lambda_{1,2} = \tfrac{a+d}{2} \pm \sqrt{\left(\tfrac{a+d}{2}\right)^2 - ad - bc}$$

$$\begin{aligned}\lambda_{1,2} &= m \pm \sqrt{h^2 + p^2 - k^2} \\ &= m \pm \Delta\end{aligned}$$

with $\Delta = \sqrt{h^2 + p^2 - k^2}$

This last formula specifically gives us some direct insight into the structure of the eigenvalues. First, the two eigenvalues are centered around m which is the trace divided by 2 or the arithmetic mean of the diagonal. (This actually extends to the $n \times n$; the arithmetic mean of the diagonal (ie. $tr(A)/n$) is the centroid of the eigenvalues.) Secondly, the geometry of the vectors

$$u_{\pm} = \begin{bmatrix} m \\ \pm k \end{bmatrix}, \qquad v_{\pm} = \begin{bmatrix} \pm h \\ p \end{bmatrix},$$

tell us a lot about the matrix and it's eigenvalues. For a symmetric matrix ($k = 0$) the eigenvalues are given by

$$\lambda_{1,2} = m \pm \sqrt{h^2 + p^2} = m \pm |v|$$

In this case, the definitiness of the matrix is determined by the relative size of m with the norm of v. A symmetric 2×2 matrix is definite if and only if

$$|m| > |v|$$

Positive or negative definite depends on the sign of m. The discrimant is given by

$$p^2 + h^2 - k^2 = |v|^2 - k^2$$

A has real eigenvalues if and only if

$$|k| \leq |v|$$

If $p = h = 0$, then the eigenvalues are given by

$$\lambda_{1,2} = m \pm ki = \sqrt{m^2 + k^2} e^{\pm i\phi}$$

with $\phi = tan^{-1}\left(\frac{k}{m}\right)$ where the second equality gives the polar form. This last characterization shows a close relationship between a matrix of the form

$$\begin{aligned} A = \begin{bmatrix} m & -k \\ k & m \end{bmatrix} &= \sqrt{m^2 + k^2} \begin{bmatrix} \frac{m}{\sqrt{m^2+k^2}} & -\frac{k}{\sqrt{m^2+k^2}} \\ \frac{k}{\sqrt{m^2+k^2}} & \frac{m}{\sqrt{m^2+k^2}} \end{bmatrix} \\ &= \sqrt{m^2 + k^2} \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix} \end{aligned}$$

and the complex conjugate pair $m \pm ki$. Indeed much of the intuition behind complex eigenvalues is grounded in understanding matrices of this form. Note that the last two equalites write the matrix as a scaled rotation closely related to the polar form of the eigenvalues. More uses for these formulas are detailed in the section on linear vector fields and stability.

B. *Eigenvectors*

While the above eigenvalue analysis is insightful, we must also consider the eigenvectors in order to get a full picture of the action of a matrix. We will discuss primarily discuss right eigenvectors here, but analogous results apply to left eigenvectors as well. Since the length of an eigenvector is irrelevant, we will present formulas for eigenvectors with the understanding that any scaling of that vector is also an eigenvector. To explicit, we will use $\cong$ to represent proportional to.

For an eigenvalue λ, the right eigenvector is contained in the nullspace of

$$\lambda I - A = \begin{bmatrix} \lambda - a & -b \\ -c & \lambda - d \end{bmatrix}$$

Row-reducing this matrix gives

$$\begin{aligned} \underbrace{\begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix}}_{E_2} \underbrace{\begin{bmatrix} \frac{1}{\lambda - a} & 0 \\ 0 & 1 \end{bmatrix}}_{E_1} \begin{bmatrix} \lambda - a & -b \\ -c & \lambda - d \end{bmatrix} &= \begin{bmatrix} 1 & -\frac{b}{\lambda - a} \\ 0 & \frac{(\lambda - d)(\lambda - a) - cb}{\lambda - a} \end{bmatrix} \\ &= \begin{bmatrix} 1 & -\frac{b}{\lambda - a} \\ 0 & 0 \end{bmatrix} \end{aligned}$$

where the last equation is from λ being a root of the characteristic polynomial. We then have the following two characterizations of a right eigenvector for λ.

$$V_{1,2} \cong \begin{bmatrix} b \\ \lambda_{1,2} - a \end{bmatrix} \cong \begin{bmatrix} \lambda_{1,2} - d \\ c \end{bmatrix}$$

The first characterization here comes from the row reduction done above (where the $1,1$ entry of the matrix is taken as the pivot); the second characterization comes from if the $2,2$ entry of the matrix is taken as the pivot. We note also that each of these vectors is clearly orthogonal to one of the rows of the matrix above Since in a rank-1 2×2 matrix the rows are just scalings of each other, being orthogonal to one row is the same as being orthogonal to other other so we could have have just read off both of these characterizations initially. (Again, note that this rank-1 condition (and thus the above eigenvector characterization) is not true for all λ but only when λ satisfies the characteristic equation.)

Along with picking vectors orthogonal to both (in this case, either) row, there is another way to read off eigenvectors based on diagonalizing $\lambda I - A$. If we diagonalize $\lambda I - A$ we get

$$\lambda I - A = \begin{bmatrix} \lambda - a & -b \\ -c & \lambda - d \end{bmatrix}$$
$$= \begin{bmatrix} | & | \\ V_1 & V_2 \\ | & | \end{bmatrix} \begin{bmatrix} \lambda - \lambda_1 & 0 \\ 0 & \lambda - \lambda_2 \end{bmatrix} \underbrace{\begin{bmatrix} - & W_1^T & - \\ - & W_2^T & - \end{bmatrix}}_{V^{-1}}$$

$$\begin{bmatrix} \lambda - a & -b \\ -c & \lambda - d \end{bmatrix} = \begin{bmatrix} | \\ V_1 \\ | \end{bmatrix} (\lambda - \lambda_1) \begin{bmatrix} - & W_1^T & - \end{bmatrix} + \begin{bmatrix} | \\ V_2 \\ | \end{bmatrix} (\lambda - \lambda_2) \begin{bmatrix} - & W_2^T & - \end{bmatrix}$$

If we plug in λ_2, then the second matrix term in the sum goes to 0 and we get that both columns of the resulting matrix are actually scalings of V_1. Similarly if we plug in λ_1, then the columns become scalings of V_2 From this we have that

$$V_1 \cong \begin{bmatrix} \lambda_2 - a \\ -c \end{bmatrix} \cong \begin{bmatrix} -b \\ \lambda_2 - d \end{bmatrix}$$

and that

$$V_2 \cong \begin{bmatrix} \lambda_1 - a \\ -c \end{bmatrix} \cong \begin{bmatrix} -b \\ \lambda_1 - d \end{bmatrix}$$

Note here that the of the eigenvalues/eigenvectors is opposite as opposed to above where it was the same. Note again that the lengths of of the eigenvectors (for each eigenvalue) here are not the same and one would need to work a little harder to show how they differ. Again each of these different subspace characterizations is only the same because λ is a root of the characteristic polynomial. Any attempt to show that these vectors have the same span will involve using the fact that $(\lambda - a)(\lambda - d) - cb = 0$. The rank of $\lambda I - A$ dropping and it's relationship to the eigenvectors is given in the following illustration

Plugging in the eigenvalues for λ to the above forms gives several specific characterizations.

$$V_{1,2} \cong \begin{bmatrix} b \\ \lambda_{1,2} - a \end{bmatrix} = \begin{bmatrix} p - k \\ -h \end{bmatrix} \pm \begin{bmatrix} 0 \\ \sqrt{p^2 + h^2 - k^2} \end{bmatrix}$$

$$V_{1,2} \cong \begin{bmatrix} \lambda_{1,2} - d \\ c \end{bmatrix} = \begin{bmatrix} h \\ p + k \end{bmatrix} \pm \begin{bmatrix} \sqrt{p^2 + h^2 - k^2} \\ 0 \end{bmatrix}$$

$$V_1 \cong \begin{bmatrix} \lambda_2 - a \\ -c \end{bmatrix} = \begin{bmatrix} -h \\ -(p + k) \end{bmatrix} - \begin{bmatrix} \sqrt{p^2 + h^2 - k^2} \\ 0 \end{bmatrix}$$

$$V_1 \cong \begin{bmatrix} -b \\ \lambda_2 - d \end{bmatrix} = \begin{bmatrix} -(p - k) \\ h \end{bmatrix} - \begin{bmatrix} 0 \\ \sqrt{p^2 + h^2 - k^2} \end{bmatrix}$$

$$V_2 \cong \begin{bmatrix} \lambda_1 - a \\ -c \end{bmatrix} = \begin{bmatrix} -h \\ -(p + k) \end{bmatrix} + \begin{bmatrix} \sqrt{h^2 + p^2 - k^2} \\ 0 \end{bmatrix}$$

$$V_2 \cong \begin{bmatrix} -b \\ \lambda_1 - d \end{bmatrix} = \begin{bmatrix} -(p - k) \\ h \end{bmatrix} + \begin{bmatrix} 0 \\ \sqrt{h^2 + p^2 - k^2} \end{bmatrix}$$

We should note that m does not appear in any of the formulas. The reason for this is that m gets added to both diagonal elements equally and thus provides the component of A strictly proportional to the identity.

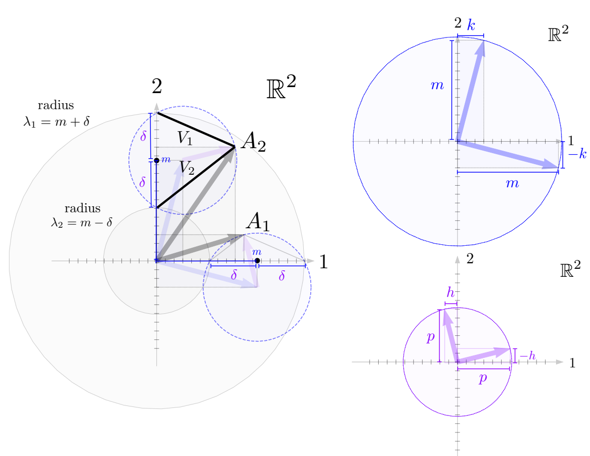
$$A = m \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} h & p - k \\ p + k & h \end{bmatrix}$$

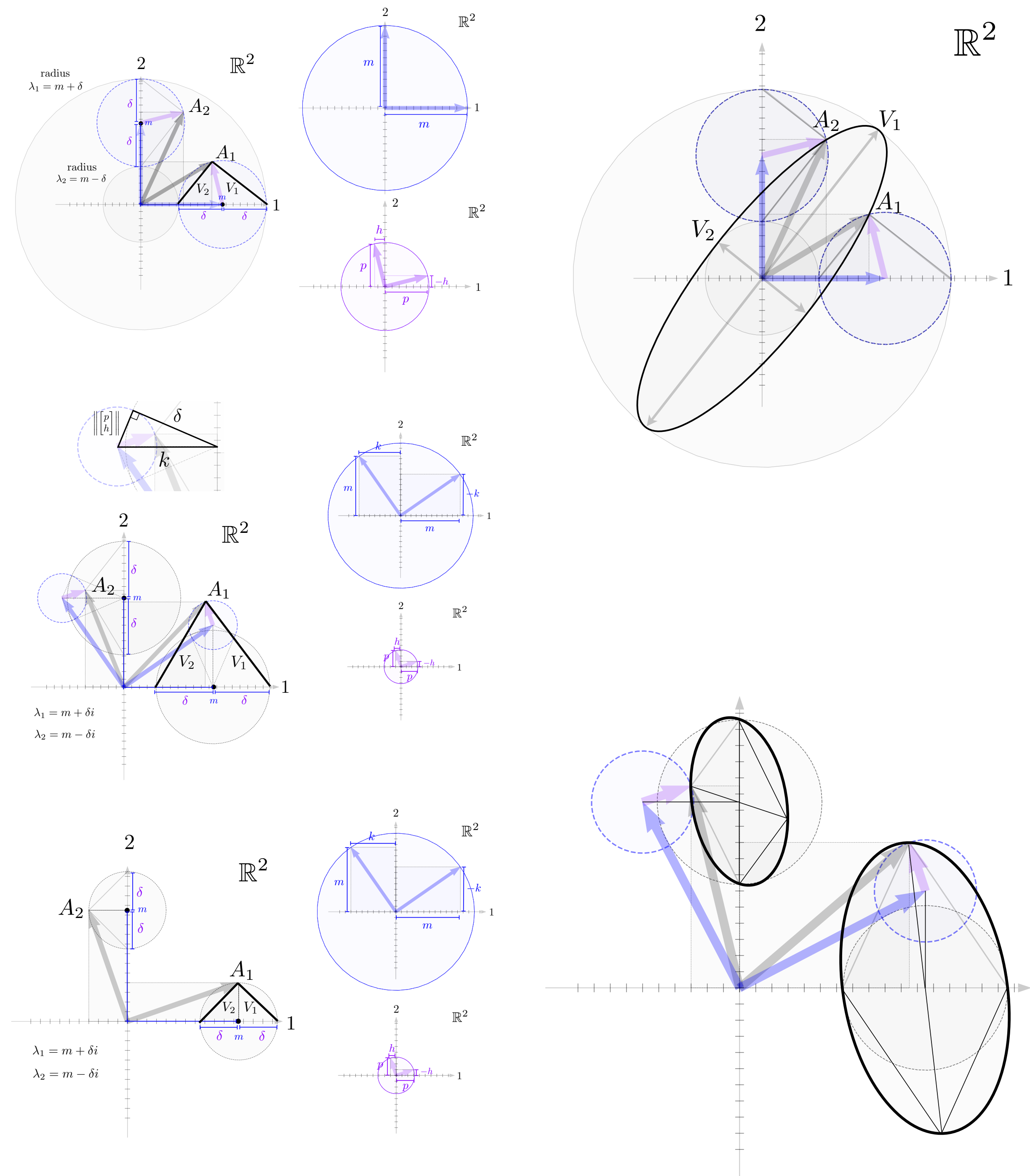
Thus for any value of m the matrix has the same eigenvectors. Since adding a scaling of the identity shifts the eigenvalues but does not change the eigenvectors this is to be expected.

C. Eigen-Structure Visualizations

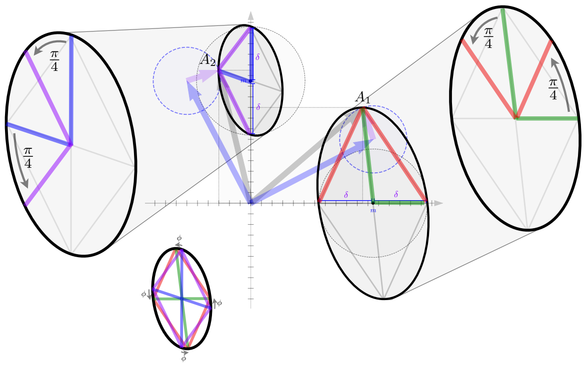
We now suggest a way to visualize the eigenstructure related to the column geometry of a 2×2 matrix. This visualization is quite dense and so we will build it up in stages.

We first look at the case where both eigenvalues are real.





For a symmetric matrix, there some interesting extensions. If we consider the image of the unit circle, we notice that the resulting ellipse elegantly inscribes/circumscribes the two eigenvalue balls and also that the the eigenvectors are orthogonal to each other.

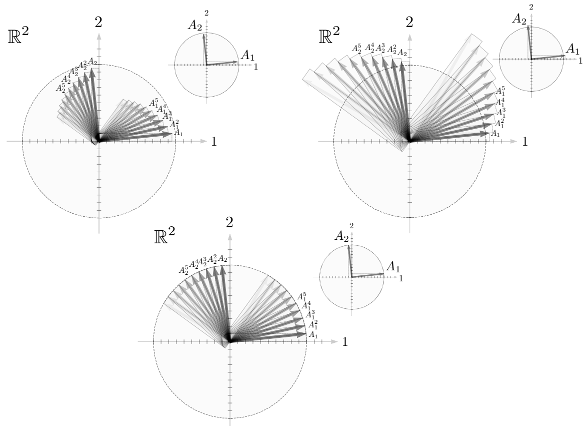


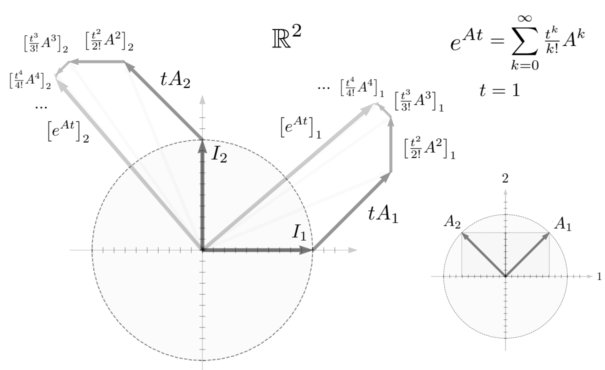
VIII. Complex Eigenvalues

Rotation shape
Rotation angle

IX. Matrix Functions

Powers of matrices Spectral mapping theorem





X. Rotation Matrices

XI. Definite Matrices

Quadratic forms
Eigenvectors (orthonormal)
Eigenvalues (real)
Positive, Negative, Indefinite

XII. Matrix Shapes

A. Grammians

When studying matrix properties, there are two positive definite matrices known as Grammians that show up frequently,

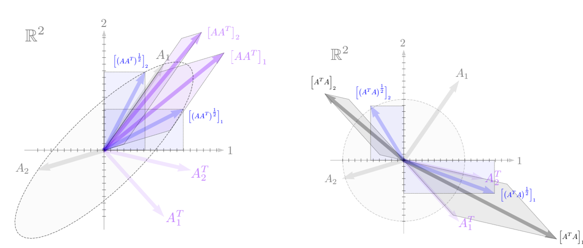
$$A^T A \in \mathbb{R}^{n\times n} \qquad AA^T \in \mathbb{R}^{m\times m}$$

for $A \in \mathbb{R}^{m\times n}$. $A^T A$ naturally arises, for example, when computing the 2-norm of a vector under a linear transformation. For $y = Ax$, $y^T y$ defines a quadratic form with matrix $A^T A$; $y^T y = x^T A^T Ax$. . It's interesting to note that these matrices are well-defined even when A is not square and that they also have the same rank as A. This comes from the rank-nullity theorem and the fact that $Ax = 0$ clearly implies $A^T Ax = 0$ and also that $A^T Ax = 0 \Rightarrow x^T A^T Ax = |Ax|^2 = 0$ which by properties of norms can only be 0 if $Ax = 0$. Similarly AA^T has the same rank as A^T (and thus A). These matrices often show up in pseudo-inverse formulas because even if A is not square if A is full row rank than AA^T is invertible and if A is full column rank than $A^T A$ is invertible. Indeed, the most common left and right pseudo-inverses are given by

$$A^L = (A^T A)^{-1} A^T, \qquad A^R = A(AA^T)^{-1}$$

It should also be noted that $A^T A$ and AA^T are both symmetric and positive semi-definite since $x^T A^T Ax = |Ax|^2 = 0$ and norms are always positive. If A is full row rank than AA^T is positive definite; if A is full column rank than $A^T A$ is positive definite.

For 2×2 matrices, we can visualize the column geometry of the matrices using the matrix transpose and matrix multiplication techniques detailed above; however, it is not clear that these visualizations are easy enough to see to be useful.



(In some ways these visualizations are best used only as a supplement to the visualizations in the next section.)

B. *Shape Matrices*

Interestingly, the values of A^TA are only determined by the relative orientation and size of the columns from each other and the values of AA^T are only determined by the relative orientation and size of the rows. Explicitly, we can see that

$$\begin{aligned} A^TA &= \begin{bmatrix} - & A_1^T & - \\ & \vdots & \\ - & A_n^T & - \end{bmatrix} \begin{bmatrix} | & & | \\ A_1 & \cdots & A_n \\ | & & | \end{bmatrix} \\ &= \begin{bmatrix} A_1^TA_1 & \cdots & A_1^TA_n \\ \vdots & & \vdots \\ A_n^TA_1 & \cdots & A_n^TA_n \end{bmatrix} \end{aligned}$$

the entries of A^TA just depend on pairwise inner producsts of the columns (and similarly with AA^T and the rows.) We can see this more compactly by considering applying a rotation $R \in \mathbb{R}^{n\times n}$ to each column of A and noticing that A^TA is unchanged.

$$(RA)^T(RA) = A^TR^TRA = A^TA$$

Even more fundamental perhaps than A^TA and AA^T are the matrix square roots of these matrices

$$(A^TA)^{\frac{1}{2}} \in \mathbb{R}^{n\times n} \qquad (AA^T)^{\frac{1}{2}} \in \mathbb{R}^{m\times m}$$

Here we choose the positive root of the eigenvalues to make these matrices positive definite. Note that any matrix square roots have the property that squaring them returns the original matrix (as was to be expected).

$$\begin{aligned} A^TA &= (A^TA)^{\frac{1}{2}}(A^TA)^{\frac{1}{2}} \\ AA^T &= (AA^T)^{\frac{1}{2}}(AA^T)^{\frac{1}{2}} \end{aligned}$$

We can refer to $(A^TA)^{\frac{1}{2}}$ as the "shape of the columns" and $(AA^T)^{\frac{1}{2}}$ as the "shape of the rows." By "shape of the columns," we mean the unique positive semi-definite matrix whose columns have the same relative shape as the columns of A. The inspiration for this comes from the relative geometry of the columns discussed above but we can explicitly check this (at least in the case where A^TA is invertible) by writing

$$A = \underbrace{A(A^TA)^{-\frac{1}{2}}}_{R}(A^TA)^{\frac{1}{2}}$$

We can then show explicitly that R is an isometry, ie. $R^TR = I$.

$$R^TR = (A^TA)^{-\frac{1}{2}}A^TA(A^TA)^{-\frac{1}{2}} = I$$

Similarly, we could write

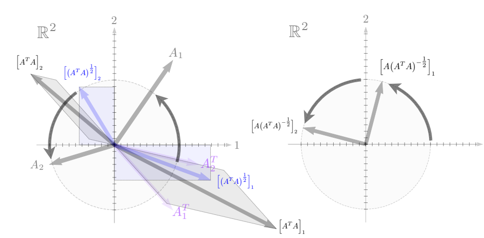
$$A = (AA^T)^{\frac{1}{2}}\underbrace{(AA^T)^{-\frac{1}{2}}A}_{R'}$$

and show that R' is an isometry to show that the rows of $(AA^T)^{\frac{1}{2}}$ are isometric to the rows of A. (It turns out for invertible A that $R = R'$ but this is not at all obvious.)

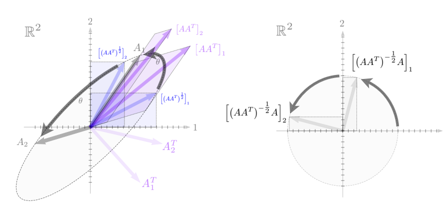
These two formulas we have just derived are known as the two polar decompositions of A, one based on the column shape, one based on the row shape.

$$\begin{aligned} A &= \underbrace{A(A^TA)^{-\frac{1}{2}}}_{\text{ortho.}}\underbrace{(A^TA)^{\frac{1}{2}}}_{\text{pos.def.}} \\ A &= \underbrace{(AA^T)^{\frac{1}{2}}}_{\text{pos. def.}}\underbrace{(AA^T)^{-\frac{1}{2}}A}_{\text{ortho.}} \end{aligned}$$

The above visualizations for A^TA and AA^T become more enlightening if we now add the matrices $(A^TA)^{\frac{1}{2}}$ and $(AA^T)^{\frac{1}{2}}$. We also draw the orthonormal transformations R and R'. The reader should note that rotating the columns of $(A^TA)^{\frac{1}{2}}$ by R moves them to A and squaring $(A^TA)^{\frac{1}{2}}$ moves it to A^TA.



Similarly, right multiplying $(AA^T)^{\frac{1}{2}}$ by R' transforms it to A and squaring it transforms it to AA^T.



(DISCUSSION OF POLAR FORM ANALOGY WITH COMPLEX NUMBERS)

The polar decompositions are precursors to the singular value decomposition in that the SVD is derived by diagonalizing A^TA or AA^T. After deriving the SVD, however, it is worth plugging the SVD matrices back into the polar decomposition formulas. We do this here (before discussing the SVD in detail for readablility). For an invertible $A \in \mathbb{R}^{n\times n}$. For SVD given by

$$A = U\Sigma V^T$$

The Grammians are given by

$$A^TA = V\Sigma^2V^T, \qquad AA^T = U\Sigma^2U^T$$

The shape matrices are given by

$$(A^TA)^{\frac{1}{2}} = V\Sigma V^T, \qquad (AA^T)^{\frac{1}{2}} = U\Sigma U^T$$

The polar decompositions are given by

$$A = R(A^TA)^{\frac{1}{2}} = \underbrace{(UV^T)}_{\text{ortho.}}\underbrace{(V\Sigma V^T)}_{\text{pos.def.}}$$
$$A = (AA^T)^{\frac{1}{2}}R' = \underbrace{(U\Sigma U^T)^{\frac{1}{2}}}_{\text{pos.def.}}\underbrace{(UV^T)}_{\text{ortho}}$$

We note from this last set of formulas that $R = R' = UV^T$ and also that A^TA and AA^T have the same non-zero eigenvalues.

For 2×2 matrices, we can the Grammian matrices in terms of the M,H decomposition. This will prove specifically useful when we diagonalize A^TA and AA^T in order to write explicit formulas for the shape matrices $(A^TA)^{\frac{1}{2}}$ and $(AA^T)^{\frac{1}{2}}$ and the singular value matrices U, Σ, V in terms of m, h, p, k (and then by proxy in terms of a, b, c, d). To this end we first give a proposition about M and H that can be checked by direct computation.

Proposition 3.

$$MH = HM^T, \quad HM = M^TH$$

Similar proofs apply to $H^{\frac{1}{2}}$ and $M^{\frac{1}{2}}$.

We then have that

$$\begin{aligned} A^TA &= (M^T + H^T)(M + H) \\ &= M^TM + H^TH + H^TM + M^TH \\ &= \big(|u|^2 + |v|^2\big)I + 2HM \end{aligned}$$

$$\begin{aligned} AA^T &= (M + H)(M^T + H^T) \\ &= MM^T + HH^T + HM^T + MH \\ &= \big(|u|^2 + |v|^2\big)I + 2MH \end{aligned}$$

where we've used the proposition above and also that the columns of M and H are orthogonal to each other.

XIII. Singular Value Decomposition

The construction of the singular value decomposition proceeds by diagonalizing one of the Grammian matrices. We will show one such derivation here for invertible A. Diagonalize A^TA as

$$A^TA = V\Sigma^2V^T$$

for orthonormal V and positive diagonal Σ^2. Another way to write this equation is that $A^TAV = V\Sigma^2$ These assumptions are valid since A^TA is symmetric positive (semi-)definite. We now notice that the columns of $AV\Sigma^{-1}$ are orthonormal eigenvectors of AA^T with eigenvalues given by the diagonal of Σ^2. Indeed

$$AA^T(AV\Sigma^{-1}) = AA^T(AV\Sigma^{-1}) = AV\Sigma^2\Sigma^{-1}) = (AV\Sigma^{-1})\Sigma^2$$

To check orthonormality, we check that $\Sigma^{-T}V^TA^TAV\Sigma^{-1} = I$. Define $U = AV\Sigma^{-1}$ and rearrange to get

$$A = U\Sigma V^T$$

Here U, V orthonormal eigenvectors of AA^T and A^TA and Σ diagonal with elements that are the square roots of the eigenvalues of the same.

From above we have that

$$\begin{aligned} A^TA &= \big(|u|^2 + |v|^2\big)I + 2HM \\ AA^T &= \big(|u|^2 + |v|^2\big)I + 2MH \end{aligned}$$

We can see from these computations that the (trace-zero matrices) MH and HM determine the eigenstructure of A^TA and AA^T and thus the singular structure of A.

A. Singular Values

The singular values are the eigenvalues of MH and HM shifted by $|u|^2 + |v|^2$.

We now write out MH and HM explicitly.

$$MH = \begin{bmatrix} m & -k \\ k & m \end{bmatrix}\begin{bmatrix} h & p \\ p & -h \end{bmatrix} = \begin{bmatrix} mh - pk & mp + kh \\ mp + kh & -mh + pk \end{bmatrix}$$
$$HM = \begin{bmatrix} h & p \\ p & -h \end{bmatrix}\begin{bmatrix} m & -k \\ k & m \end{bmatrix} = \begin{bmatrix} mh + kp & mp - kh \\ mp - kh & -mh - kp \end{bmatrix}$$

We can compute that the eigenvalues of MH and HM

are given by

$$\begin{aligned}\lambda_{1,2}(MH) &= \pm\sqrt{(mh-kp)^2+(mp+hk)^2}\\ &= \pm\sqrt{m^2h^2+k^2p^2+m^2p^2+h^2k^2}\\ &= \pm\sqrt{(m^2+k^2)(p^2+h^2)}\\ &= \pm|u||v|\\ \lambda_{1,2}(HM) &= \pm\sqrt{(mh+kp)^2+(mp-hk)^2}\\ &= \pm\sqrt{m^2h^2+k^2p^2+m^2p^2+h^2k^2}\\ &= \pm\sqrt{(m^2+k^2)(p^2+h^2)}\\ &= \pm|u||v|\end{aligned}$$

Note these eigenvalues are the same as expected from the fact that A^TA and AA^T have the same eigenvalues. We then have that the eigenvalues of A^TA and AA^T are given by

$$\begin{aligned}\lambda_{1,2}(A^TA) = \lambda_{1,2}(AA^T) &= |u|^2+|v|^2 \pm 2|u||v|\\ &= (|u| \pm |v|)^2\end{aligned}$$

Finally we then have that the singular values of A are given by

$$\begin{aligned}\sigma_{1,2} &= \Big||u| \pm |v|\Big|\\ &= \left|\sqrt{m^2+k^2} \pm \sqrt{p^2+h^2}\right|\end{aligned}$$

We note this formula is remarkably elegant and clean and geometrically can be shown as follows.

B. Singular Vectors

We now turn our attention to the singular vectors, ie. the eigenvectors of AA^T, $(AA^T)^{\frac{1}{2}}$ and of A^TA, $(A^TA)^{\frac{1}{2}}$. The most direct path forward is to consider the eigenstructures of MH and HM. We include a longer derivation of the singular vectors following this method in the appendix. Here, however, we give a quite simple derivation that is much easier to follow though definitely harder to find.

We will require a simple lemma about 2×2 rotation matrices.

Lemma 1. *For a 2×2 rotation matrix*

$$\begin{bmatrix}\cos\phi & -\sin\phi\\ \sin\phi & \cos\phi\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}\cos\phi & \sin\phi\\ -\sin\phi & \cos\phi\end{bmatrix}$$

Our proof comes from direct analysis of M and H which we reproduce here for clarity.

$$M = \begin{bmatrix}m & -k\\ k & m\end{bmatrix}, \qquad H = \begin{bmatrix}h & p\\ p & -h\end{bmatrix},$$

Define angles ϕ and ψ given by

$$\phi = \arctan\left(\frac{k}{m}\right), \qquad \psi = \arctan\left(\frac{p}{h}\right)$$

Here ϕ is the rotation angle of u up from the first axis and ψ is the rotation of v from the first axis. For any angle θ let R_θ denote a rotation matrix by θ . Note that

$$M = |u|R_\phi, \qquad H = |v|R_\psi\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$$

Define two rotations

$$U = R_{\frac{\phi+\psi}{2}} = R_{\phi/2}R_{\psi/2} = \begin{bmatrix}\cos\left(\frac{\phi+\psi}{2}\right) & -\sin\left(\frac{\phi+\psi}{2}\right)\\ \sin\left(\frac{\phi+\psi}{2}\right) & \cos\left(\frac{\phi+\psi}{2}\right)\end{bmatrix}$$

$$V = R_{\frac{\psi-\phi}{2}} = R^T_{\phi/2}R_{\psi/2} = \begin{bmatrix}\cos\left(\frac{\psi-\phi}{2}\right) & -\sin\left(\frac{\psi-\phi}{2}\right)\\ \sin\left(\frac{\psi-\phi}{2}\right) & \cos\left(\frac{\psi-\phi}{2}\right)\end{bmatrix}$$

Here U is the rotation caused by the average of the two phase angles phi and ψ and V is the rotation caused by the offset phase. The notation U and V is also intentional since we will show that they define the singular vectors. More explicit formulas for U and V can be derived via trig identities as

$$U = \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix}\sqrt{1+\frac{m}{|u|}} & -\sqrt{1-\frac{m}{|u|}}\\ \sqrt{1-\frac{m}{|u|}} & \sqrt{1+\frac{m}{|u|}}\end{bmatrix}}_{R_{\phi/2}}\underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix}\sqrt{1+\frac{h}{|v|}} & -\sqrt{1-\frac{h}{|v|}}\\ \sqrt{1-\frac{h}{|v|}} & \sqrt{1+\frac{h}{|v|}}\end{bmatrix}}_{R_{\psi/2}}$$

$$V = \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix}\sqrt{1+\frac{m}{|u|}} & \sqrt{1-\frac{m}{|u|}}\\ -\sqrt{1-\frac{m}{|u|}} & \sqrt{1+\frac{m}{|u|}}\end{bmatrix}}_{R^T_{\phi/2}}\underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix}\sqrt{1+\frac{h}{|v|}} & -\sqrt{1-\frac{h}{|v|}}\\ \sqrt{1-\frac{h}{|v|}} & \sqrt{1+\frac{h}{|v|}}\end{bmatrix}}_{R^T_{\psi/2}}$$

Note also that M and H can be written in terms of U and V as well. The formula for M is straightforward.

$$M = |u|R_\phi = |u|R_{\phi/2}R_{\psi/2}R^T_{\psi/2}R_{\phi/2} = |u|UV^T$$

The formula for H requires applications of the lemma above.

$$\begin{aligned}H &= |v|R_\psi\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\\ &= |v|R_{\psi/2}R_{\phi/2}R^T_{\phi/2}R_{\psi/2}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\\ &= |v|UV\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\\ &= |v|U\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}V^T\end{aligned}$$

Finally, now we can construct the SVD directly.

$$
\begin{aligned}
A &= M + H \\
&= |u| U V^T + |v| U \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} V^T \\
&= U \begin{bmatrix} |u| + |v| & 0 \\ 0 & |u| - |v| \end{bmatrix} V^T
\end{aligned}
$$

We note that this is basically the SVD except when the case where $|v| > |u|$. Here we simply multiply the second column of U (or V) by -1 and replace $|u| - |v|$ with $||u| - |v||$.

In general then, we can take

$$
\Sigma = \begin{bmatrix} |u| + |v| & 0 \\ 0 & ||u| - |v|| \end{bmatrix}
$$

and use the formulas for U and V given above with the understanding that we flip the sign on one column of U or V if needed (when $|v| > |u|$).

C. Explicit Formulas for Polar Decomposition

These derivations can be applied to the polar decomposition matrices to acquire specific formulas for them as well. The Grammian matrices $A^T A$ and AA^T are quite easy to write explicitly. The polar rotation can be constructed as

$$
UV^T = R_\phi = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} = \begin{bmatrix} \frac{m}{|u|} & -\frac{k}{|u|} \\ \frac{k}{|u|} & \frac{m}{|u|} \end{bmatrix}
$$

Perhaps surprisingly, this is just M with normalized columns. To construct the shape matrices $(A^T A)^{\frac{1}{2}}$ and $(AA^T)^{\frac{1}{2}}$ we first give the following lemma that details 2×2 symmetric matrix structure. For a diagonalizable symmetric 2×2 matrix with eigenvectors defined by the columns of rotation matrix R we have

Lemma 2. *Symmetric matrix structure*

$$
\begin{aligned}
RDR^T &= \begin{bmatrix} c\phi & -s\phi \\ s\phi & c\phi \end{bmatrix} \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \begin{bmatrix} c\phi & s\phi \\ -s\phi & c\phi \end{bmatrix} \\
&= \begin{bmatrix} c\phi & -s\phi \\ s\phi & c\phi \end{bmatrix} \begin{bmatrix} \alpha + \beta & 0 \\ 0 & \alpha - \beta \end{bmatrix} \begin{bmatrix} c\phi & s\phi \\ -s\phi & c\phi \end{bmatrix} \\
&= \alpha \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \beta \begin{bmatrix} c2\phi & s2\phi \\ s2\phi & -c2\phi \end{bmatrix}
\end{aligned}
$$

Applying this lemma to $(A^T A)^{\frac{1}{2}} = V \Sigma V^T$ and $(AA^T)^{\frac{1}{2}} = U \Sigma U^T$ gives us

$$
\begin{aligned}
(A^T A)^{\frac{1}{2}} &= V \Sigma V^T = \alpha I + \beta \begin{bmatrix} \mathrm{c}\,(\psi - \phi) & \mathrm{s}\,(\psi - \phi) \\ \mathrm{s}\,(\psi - \phi) & -\mathrm{c}\,(\psi - \phi) \end{bmatrix} \\
(AA^T)^{\frac{1}{2}} &= U \Sigma U^T = \alpha I + \beta \begin{bmatrix} \mathrm{c}\,(\phi + \psi) & \mathrm{s}\,(\phi + \psi) \\ \mathrm{s}\,(\phi + \psi) & -\mathrm{c}\,(\phi + \psi) \end{bmatrix} \\
& \alpha = \max(|u|, |v|), \quad \beta = \min(|u|, |v|)
\end{aligned}
$$

We note that there are two forms of each of these matrices depending on whether $|u| > |v|$ or vice versa.

Alternatively these equations could be written

$$
\begin{aligned}
(A^T A)^{\frac{1}{2}} &= \alpha I + \beta V^2 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\
(AA^T)^{\frac{1}{2}} &= \alpha I + \beta U^2 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\
\alpha &= \max(|u|, |v|), \quad \beta = \min(|u|, |v|)
\end{aligned}
$$

which is again perhaps surprising.

XIV. Linear Vector Fields

From the above analysis, it seems profitable to plot $u_\pm$ and $v_\pm$ in a 2D space analogous to the complex plane. For $v_\pm = 0$ this space is precisely a picture of the complex plane and the vectors $u_\pm$ are the eigenvalues of the matrix. When $v_\pm \neq 0$ we can modify the picture in the following way. Plot the vector $v_\pm$ and the ball it touches. Properties of the eigenvalues of A are then given by what region $u_\pm$ falls in relative to the ball generated by $v_\pm$. These regions are shown in the diagram below

XV. Appendices

A. Eigenvector characterizations

B. Singular Decomposition (More Details)

NOTES TO SELF:

$$
\begin{aligned}
A^T A &= \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \\
&= \begin{bmatrix} a^2 + c^2 & ab + dc \\ ab + cd & b^2 + d^2 \end{bmatrix}
\end{aligned}
$$

Proposition 4.

$$
\begin{bmatrix} m & -k \\ k & m \end{bmatrix}^{\frac{1}{2}} = \sqrt{|u|} \underbrace{\begin{bmatrix} \cos\frac{\phi}{2} & -\sin\frac{\phi}{2} \\ \sin\frac{\phi}{2} & \cos\frac{\phi}{2} \end{bmatrix}}_{R_{\phi/2}} = \frac{1}{\sqrt{2}} \begin{bmatrix} \sqrt{|u| + m} & -\sqrt{|u| - m} \\ \sqrt{|u| - m} & \sqrt{|u| + m} \end{bmatrix}
$$

Proof: Direct computation. Explicitly:

$$
\begin{aligned}
M^{\frac{1}{2}} M^{\frac{1}{2}} &= \frac{1}{2} \begin{bmatrix} \sqrt{|u| + m} & -\sqrt{|u| - m} \\ \sqrt{|u| - m} & \sqrt{|u| + m} \end{bmatrix} \begin{bmatrix} \sqrt{|u| + m} & -\sqrt{|u| - m} \\ \sqrt{|u| - m} & \sqrt{|u| + m} \end{bmatrix} \\
&= \frac{1}{2} \begin{bmatrix} 2m & -2\sqrt{(|u| + m)(|u| - m)} \\ 2\sqrt{(|u| + m)(|u| - m)} & 2m \end{bmatrix} \\
&= \frac{1}{2} \begin{bmatrix} 2m & -2\sqrt{|u|^2 - m^2} \\ 2\sqrt{|u|^2 - m^2} & 2m \end{bmatrix} \\
&= \begin{bmatrix} m & -\sqrt{m^2 + k^2 - m^2} \\ \sqrt{m^2 + k^2 - m^2} & m \end{bmatrix} \\
&= \begin{bmatrix} m & -k \\ k & m \end{bmatrix}
\end{aligned}
$$

$\blacksquare$

NOTE TO SELF: For HM: diagonal angle is $\phi - \psi$ and off diagonal angle is $\phi + \frac{\pi}{2} - \psi$ where ϕ is the rotation angle of M and ψ is the rotation angle of H.

$$\phi - \psi, \quad \phi + \tfrac{\pi}{2} - \psi$$

For MH: (above just replace ϕ with $-\phi$)

$$-\phi - \psi, \quad -\phi + \tfrac{\pi}{2} - \psi$$

$$\begin{bmatrix} mh - kp & mp + hk \\ mp + hk & -(mh - kp) \end{bmatrix} - \begin{bmatrix} -|u||v| & 0 \\ 0 & |u||v| \end{bmatrix}$$

$$|u||v| \begin{bmatrix} \Big(1 + \cos\Big(-\phi - \psi\Big)\Big) & \cos\Big(-\phi - \psi + \frac{\pi}{2}\Big) \\ \cos\Big(-\phi - \psi + \frac{\pi}{2}\Big) & -\Big(1 + \cos\Big(-\phi - \psi\Big)\Big) \end{bmatrix}$$

Dividing out the factor of $|u||v|$, multiplying the second column by -1, and applying several trig identities gives

$$\begin{bmatrix} 1 + \cos\Big(-\phi - \psi\Big) & -\sin\Big(-\phi - \psi\Big) \\ \sin\Big(-\phi - \psi\Big) & 1 + \cos\Big(-\phi - \psi\Big) \end{bmatrix}$$

We note that the norm of each column is given by

$$\begin{aligned} \sqrt{(1 + \cos)^2 + \sin^2} &= \sqrt{1 + 2\cos + \cos^2 + \sin^2} \\ &= \sqrt{2(1 + \cos)} \end{aligned}$$

Normalizing each column gives

$$\begin{bmatrix} \sqrt{\frac{1}{2}(1 + \cos)} & -\sqrt{\frac{1}{2}(1 + \cos)}\frac{\sin}{1+\cos} \\ \sqrt{\frac{1}{2}(1 + \cos)}\frac{\sin}{1+\cos} & \sqrt{\frac{1}{2}(1 + \cos)} \end{bmatrix}$$

And applying the half angle identities

$$\cos\left(\frac{\theta}{2}\right) = \pm\sqrt{\frac{1 + \cos\theta}{2}}, \quad \tan\left(\frac{\theta}{2}\right) = \frac{\sin\theta}{1 + \cos\theta}$$

gives

$$\begin{bmatrix} \cos\frac{-\phi-\psi}{2} & \cos\frac{-\phi-\psi}{2}\frac{\sin\frac{-\phi-\psi}{2}}{\cos\frac{-\phi-\psi}{2}} \\ -\cos\frac{-\phi-\psi}{2}\frac{\sin\frac{-\phi-\psi}{2}}{\cos\frac{-\phi-\psi}{2}} & \cos\frac{-\phi-\psi}{2} \end{bmatrix}$$

and finally simplifying

$$U = \begin{bmatrix} \cos\frac{-\phi-\psi}{2} & \sin\frac{-\phi-\psi}{2} \\ -\sin\frac{-\phi-\psi}{2} & \cos\frac{-\phi-\psi}{2} \end{bmatrix} = \begin{bmatrix} \cos\frac{\phi+\psi}{2} & -\sin\frac{\phi+\psi}{2} \\ \sin\frac{\phi+\psi}{2} & \cos\frac{\phi+\psi}{2} \end{bmatrix}$$

These are the eigenvectors of MH, the eigenvectors of AA^T and $(AA^T)^{\frac{1}{2}}$, and the singular vectors of U. The angles here ϕ and ψ are given by

$$\begin{aligned} \phi &= \arctan\left(\frac{k}{m}\right) \\ \psi &= \arctan\left(\frac{p}{h}\right) \end{aligned}$$

Using the following identities

$$\begin{aligned} \arctan\alpha + \arctan\beta &= \arctan\left(\frac{\alpha + \beta}{1 - \alpha\beta}\right) \\ \arctan\alpha - \arctan\beta &= \arctan\left(\frac{\alpha - \beta}{1 + \alpha\beta}\right) \end{aligned}$$

which gives that

$$\begin{aligned} \tfrac{1}{2}(\phi + \psi) &= \tfrac{1}{2}\arctan\left(\frac{k/m + p/h}{1 - (kp)/(mh)}\right) \\ &= \tfrac{1}{2}\arctan\left(\frac{mp + kh}{mh - kp}\right) \\ \tfrac{1}{2}(\phi - \psi) &= \tfrac{1}{2}\arctan\left(\frac{k/m - p/h}{1 + (kp)/(mh)}\right) \\ &= \tfrac{1}{2}\arctan\left(\frac{mp - kh}{mh + kp}\right) \end{aligned}$$

and the identities give that

$$\begin{aligned} \cos\arctan x &= \frac{1}{\sqrt{1 + x^2}} \\ \sin\arctan x &= \frac{x}{\sqrt{1 + x^2}} \end{aligned}$$

gives that

$$\begin{aligned} \cos(\phi + \psi) &= \frac{1}{\sqrt{1 + \frac{mp+kh}{mh-kp}}} \\ \sin(\phi + \psi) &= \frac{\frac{mp+kh}{mh-kp}}{\sqrt{1 + \frac{mp+kh}{mh-kp}}} \\ \cos(\phi - \psi) &= \frac{1}{\sqrt{1 + \frac{mp-kh}{mh+kp}}} \\ \sin(\phi - \psi) &= \frac{\frac{mp-kh}{mh+kp}}{\sqrt{1 + \frac{mp-kh}{mh+kp}}} \end{aligned}$$

Reorganizing gives

$$\cos\left(\phi+\psi\right) = \frac{1}{\sqrt{1+\frac{mp+kh}{mh-kp}}}$$
$$\sin\left(\phi+\psi\right) = \frac{\frac{mp+kh}{mh-kp}}{\sqrt{1+\frac{mp+kh}{mh-kp}}}$$
$$\cos\left(\phi-\psi\right) = \frac{1}{\sqrt{1+\frac{mp-kh}{mh+kp}}}$$
$$\sin\left(\phi-\psi\right) = \frac{\frac{mp-kh}{mh+kp}}{\sqrt{1+\frac{mp-kh}{mh+kp}}}$$

$$\cos\frac{\phi+\psi}{2} =$$
$$\sin\frac{\phi+\psi}{2} =$$
$$\cos\frac{\phi-\psi}{2} =$$
$$\sin\frac{\phi-\psi}{2} =$$

Replacing ϕ with $-\phi$ and applying a similar argument gives that the eigenvectors of HM are given by

$$V = \begin{bmatrix} \cos\frac{\phi-\psi}{2} & -\sin\frac{\phi-\psi}{2} \\ \sin\frac{\phi-\psi}{2} & \cos\frac{\phi-\psi}{2} \end{bmatrix} = \begin{bmatrix} \cos\frac{\psi-\phi}{2} & \sin\frac{\psi-\phi}{2} \\ -\sin\frac{\psi-\phi}{2} & \cos\frac{\psi-\phi}{2} \end{bmatrix}$$

These are the eigenvectors of HM, the eigenvectors of A^TA and $(A^TA)^{\frac{1}{2}}$, and the singular vectors of V.