Linear Algebra for Theoretical Neuroscience (Part 3)

Ken Miller

©2001, 2008 by Kenneth Miller. This work is licensed under the Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/us/ or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

Current versions of all parts of this work can be found at http://www.neurotheory.columbia.edu/∼ken/math-notes. Please feel free to link to this site.

Note: This section (Part 3) is out of date. While everything in it is correct and worth understanding, it misses a major point, that non-normal matrices can generate interesting dynamics that are not predicted by the eigenvalues. I hope to rewrite this section accordingly. A lovely book on non-normal matrices is L.N. Trefethen and M. Embree, Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators. Princeton University Press, 2005.

5 Linear Algebra and Linear Differential Equations for General Matrices

Until now, we have restricted attention to matrices with complete orthonormal sets of eigenvectors. This means we restricted to the set of *normal* matrices, meaning those that commute with their adjoint, $\mathbf{N}\mathbf{N}^\dagger = \mathbf{N}^\dagger\mathbf{N}$. This also restricted us to square matrices (normal matrices are necessarily square matrices, because if $\mathbf{N}$ is a non-square rectangular matrix then $\mathbf{N}\mathbf{N}^\dagger$ and $\mathbf{N}^\dagger\mathbf{N}$ are each square matrices with different numbers of dimensions and so cannot be equal to one another; *e.g.* if $\mathbf{N}$ is 4×6 then $\mathbf{N}\mathbf{N}^\dagger$ is 4×4 while $\mathbf{N}^\dagger\mathbf{N}$ is 6×6 – check this!). We also restricted attention to orthogonal or unitary changes of basis, since these were sufficient to get to the eigenvector basis if eigenvectors were orthonormal.

We now consider general matrices and general linear changes of basis. The general case can be understood almost as simply as the orthonormal case, and represents just a slight generalization of the formalisms we've already learned.

5.1 Some Basics: Subspaces, Bases, and Dimension

We have been using the idea of an N-dimensional vector space very loosely. It was intuitively obvious that N orthonormal vectors form a basis for an N-dimensional space, or that all linear combinations of $k < N$ orthonormal vectors form a k-dimensional subspace. Now we have to decide when *any* collection of vectors forms a basis for a space or a subspace, and how many dimensions that space or subspace has. To do this, we need to be a little more precise than we have been. We will find that a key concept is the **independence** of a set of vectors, which will be defined below: any N *independent* vectors in an N-dimensional space form a basis for that space, and given k *independent* vectors, the set of all linear combinations of them form a k-dimensional subspace.

We start with a set of scalars, either $\mathcal{R}$, the set of real numbers, or $\mathcal{C}$, the set of complex numbers. N-dimensional vectors are defined essentially as N-tuples of scalars, that is as elements of $\mathcal{R}^N$ or $\mathcal{C}^N$, along with the definition of addition of vectors (two vectors add by adding their components) and multiplication of vectors by a scalar (defined as multiplication of each component by the scalar). In what follows, we will simply refer to scalars, which should be taken as complex numbers if dealing with complex vectors or real numbers if restricting to real vectors.

We will outline the basic definitions needed to establish the concepts of subspaces, bases, and the dimension of a subspace. Along the way we will quote, mostly without proof, a few theorems that should be intuitively obvious.

Definition 5.1 *A set $\mathcal{F}$ of N-dimensional vectors is called a* **subspace** *(of $\mathcal{R}^N$ or $\mathcal{C}^N$) if it is closed under addition and scalar multiplication, that is, if for all $\mathbf{v} \in \mathcal{F}$, $\mathbf{w} \in \mathcal{F}$, and scalar λ, $\mathbf{v} + \mathbf{w} \in \mathcal{F}$ and $\lambda\mathbf{v} \in \mathcal{F}$.*

Note that a subspace might be the entire space $\mathcal{R}^N$ or $\mathcal{C}^N$, or it might be only a part of the space. We shall use the term **vector space** to refer to any subspace of $\mathcal{R}^N$ or $\mathcal{C}^N$.

Definition 5.2 *Given a set of N-dimensional vectors $\mathbf{v}_i$, $0 = 1, \ldots, k-1$, a* **linear combination** *of the $\mathbf{v}_i$ is a vector $\sum_i c_i\mathbf{v}_i$ for some set of scalars c_i. We will call it a* **non-trivial linear combination** *if at least one c_i is nonzero.*

It should be obvious from the above that, if $\mathbf{v}_i$, $i = 0, \ldots, k-1$ is some set of vectors in a vector space $\mathcal{F}$, then every linear combination of the $\mathbf{v}_i$ is also in $\mathcal{F}$ (if it's not obvious, stop and prove it!).

Definition 5.3 *Given a set of N-dimensional vectors $\mathbf{v}_i$, $i = 0, \ldots, k-1$, the* **subspace spanned by the $\mathbf{v}_i$** *is the set of all vectors that can be obtained as linear combinations of the $\mathbf{v}_i$. (If it's not obvious, show that this is indeed a subspace.)*

Definition 5.4 *For any vector space $\mathcal{F}$, a set of vectors $\mathbf{v}_i \in \mathcal{F}$, $i = 0, \ldots, k-1$, is said to* **span** *$\mathcal{F}$ if every vector in $\mathcal{F}$ can be written as a linear combination of the $\mathbf{v}_i$.*

Definition 5.5 *A set of N-dimensional vectors $\mathbf{v}_i$, $i = 0, \ldots, k-1$ is* **independent** *if no non-trivial linear combination of them is zero.*

The idea of independence is intuitively simple: if the $\mathbf{v}_i$ are not independent, then there are some c_i, not all zero, such that $\sum_i c_i\mathbf{v}_i = 0$. So suppose $c_j \neq 0$, then $\mathbf{v}_j = \sum_{i \neq j} c_i\mathbf{v}_i/c_j$, so $\mathbf{v}_j$ can be constructed as a linear combination of the other vectors. As a result, the space spanned by the $\mathbf{v}_i$ for $i \neq j$ is the same as the space spanned by all of the $\mathbf{v}_i$ – adding $\mathbf{v}_j$ to the others doesn't add anything. On the other hand, if the $\mathbf{v}_i$ are independent, then each $\mathbf{v}_i$ has some component that cannot be reached as any linear combination of the other vectors – each one contributes something unique.

Definition 5.6 *A* **basis** *of a vector space $\mathcal{F}$ is a set of vectors in $\mathcal{F}$ that (1) are independent and (2) span $\mathcal{F}$.*

Theorem 5.1 *Every vector space $\mathcal{F}$ has a basis, and every basis of $\mathcal{F}$ has the same number of elements.*

Definition 5.7 *The number of elements in a basis of $\mathcal{F}$ is called the* **dimension** *of $\mathcal{F}$.*

Theorem 5.2 *If D is the dimension of $\mathcal{F}$, then any set of D independent vectors in $\mathcal{F}$ is a basis of $\mathcal{F}$.*

Theorem 5.3 *If the elements of $\mathcal{F}$ are N-dimensional vectors, then the dimension of $\mathcal{F}$ is $\leq N$ (one cannot have a set of more than N independent N-dimensional vectors).*

Theorem 5.4 *If $\mathbf{e}_i$, $i = 0, \ldots, D-1$ is a basis for $\mathcal{F}$, then every vector in $\mathcal{F}$ can be expressed as a* **unique** *linear combination of the $\mathbf{e}_i$.*

Proof: For some $\mathbf{v} \in \mathcal{F}$, suppose $\mathbf{v} = \sum_i c_i \mathbf{e}_i$ and $\mathbf{v} = \sum_i d_i \mathbf{e}_i$. Then $0 = \mathbf{v} - \mathbf{v} = \sum_i (c_i - d_i)\mathbf{e}_i$. But the $\mathbf{e}_i$ are independent, so $c_i - d_i = 0$ for all i.

So at this point we have established our basic results: any k independent vectors in a k-dimensional subspace form a basis for that subspace, and every vector in the subspace can be represented as a unique linear combination of the basis vectors. It will also prove useful to understand the set of vectors orthogonal to a subspace:

Definition 5.8 *We say a vector* $\mathbf{v}$ *is* **orthogonal** *to a subspace $\mathcal{F}$ if it is orthogonal to every vector in $\mathcal{F}$: $\mathbf{v} \cdot \mathbf{w} = 0$ for every $\mathbf{w} \in \mathcal{F}$. We say two subspaces $\mathcal{F}_1$ and $\mathcal{F}_2$ are* **orthogonal subspaces** *if every vector in one is orthogonal to every vector in the other.*

Theorem 5.5 *If $\mathcal{F}$ is a D-dimensional subspace of $\mathcal{C}^N$ or $\mathcal{R}^N$, then the set of vectors orthogonal to $\mathcal{F}$ forms an $N - D$-dimensional subspace which is orthogonal to $\mathcal{F}$.*

5.2 Nullspace, Range, and Rank of a Matrix

Let $\mathcal{S}$ stand for either $\mathcal{R}$ or $\mathcal{C}$. Consider a matrix $\mathbf{M}$ mapping vectors in $\mathcal{S}^N$ to vectors in $\mathcal{S}^P$, that is, $\mathbf{M}$ is a $P \times N$ matrix. We refer to $\mathcal{S}^N$ as the **domain** of $\mathbf{M}$, the space upon which it acts.

Definition 5.9 *The* **nullspace** *of a $P \times N$ matrix $\mathbf{M}$ is the subspace of $\mathcal{S}^N$ that $\mathbf{M}$ maps to 0, that is, it is $\{\mathbf{v} \in \mathcal{S}^N : \mathbf{M}\mathbf{v} = 0\}$. (Prove that this is a subspace.)*

Definition 5.10 *The* **rowspace** *of a $P \times N$ matrix $\mathbf{M}$ is the subspace of $\mathcal{S}^N$ spanned by the rows of $\mathbf{M}$.*

Each row of $\mathbf{M}$ can be thought of as an N-dimensional vector (strictly, we are referring to the column vector given by the adjoint of the given row vector; but we will just call this the row). The nullspace of $\mathbf{M}$ is the set of vectors that are orthogonal to every row of $\mathbf{M}$: the k^{th} element of $\mathbf{M}\mathbf{v}$ is the dot product of the k^{th} row of $\mathbf{M}$ with $\mathbf{v}$, and each such dot product must be zero if $\mathbf{M}\mathbf{v} = 0$. If the rowspace has dimension D, then the set of vectors orthogonal to this subspace has dimension $N - D$, so:

Theorem 5.6 *Let D_r be the dimension of the rowspace of $\mathbf{M}$. Then the nullspace of $\mathbf{M}$ has dimension $N - D_r$. The rowspace and the nullspace are orthogonal subspaces.*

Definition 5.11 *The* **range** *of a $P \times N$ matrix $\mathbf{M}$ is the subspace of $\mathcal{S}^P$ that can be reached by the action of $\mathbf{M}$ on $\mathcal{S}^N$, that is, it is $\{\mathbf{v} \in \mathcal{S}^P : \mathbf{v} = \mathbf{M}\mathbf{w}$ for some $\mathbf{w} \in \mathcal{S}^N\}$. (Prove that this is a subspace.)*

Definition 5.12 *The* **column space** *of a $P \times N$ matrix $\mathbf{M}$ is the subspace of $\mathcal{S}^P$ spanned by the columns of $\mathbf{M}$.*

Each column of $\mathbf{M}$ can be thought of as a P-dimensional vector. The range of $\mathbf{M}$ is precisely the column space of $\mathbf{M}$, as can be seen as follows. Let $\mathbf{c}_j$, $j = 1, \ldots, P$ be the columns of $\mathbf{M}$; then for any vector $\mathbf{v}$, $\mathbf{M}\mathbf{v} = \sum_j v_j \mathbf{c}_j$. (This was shown in Problem 3; recall that it can be seen by writing $(\mathbf{M}\mathbf{v})_i = \sum_j M_{ij} v_j = \sum_j v_j (\mathbf{c}_j)_i$). So:

Theorem 5.7 *The range of* $\mathbf{M}$ *is equal to the column space of* $\mathbf{M}$*. In particular, if* D_c *is the dimension of the column space of* $\mathbf{M}$*, then the range of* $\mathbf{M}$ *also has dimension* D_c*.*

This brings us to what some have called the **Fundamental Theorem of Linear Algebra**. Let $\dim(\mathcal{F})$ be the dimension of a subspace $\mathcal{F}$. Then:

Theorem 5.8 *Let* $\mathbf{M} : \mathcal{S}^N \to \mathcal{S}^P$ *be a* $P \times N$ *matrix. Then dim(range* $\mathbf{M}$*) + dim(nullspace* $\mathbf{M}$*) =* N*. Equivalently, the dimension of the subspace of* $\mathcal{S}^N$ *spanned by the rows of* $\mathbf{M}$ *is equal to the dimension of the subspace of* $\mathcal{S}^P$ *spanned by the columns of* $\mathbf{M}$*:* $D_r = D_c$*.*

Thus, the number of independent rows of $\mathbf{M}$ is equal to the number of independent columns of $\mathbf{M}$. This theorem should make some intuitive sense: the elements of the nullspace don't add to the range of $\mathbf{M}$, because if $\mathbf{v}$ is in the nullspace of $\mathbf{M}$ ($\mathbf{Mv} = 0$), then for every vector $\mathbf{w}$, $\mathbf{M}(\mathbf{w} + \mathbf{v}) = \mathbf{Mw}$. So the dimension of the range is reduced, compared to the dimension N of the domain, by the dimension of the nullspace.

In particular, suppose $\mathbf{M}$ is a square matrix. If it has a nonzero nullspace, say $\mathbf{Mv} = 0$ for $\mathbf{v} \neq 0$, then we cannot "undo the mapping": if $\mathbf{y} = \mathbf{Mw}$, then also $\mathbf{y} = \mathbf{M}(\mathbf{w} + k\mathbf{v})$ for any scalar k, so we cannot compute $\mathbf{M}^{-1}\mathbf{y}$ since many different vectors map to $\mathbf{y}$. (We could, though, compute the "pseudo-inverse", by specifying that we map back only to vectors orthogonal to the nullspace – for each $\mathbf{y}$ in the range of $\mathbf{M}$, there is only one vector orthogonal to the nullspace that maps to $\mathbf{y}$.) Conversely, if no nonzero vector maps to zero, then the mapping maps N dimensions to N dimensions and it is 1-1 and can be inverted. Of course, if $\mathbf{Mv} = 0$ for $\mathbf{v} \neq 0$, then $\mathbf{v}$ is an eigenvector of $\mathbf{M}$ with eigenvalue 0, so $\mathbf{M}$ has a zero eigenvalue. So we have motivated:

Theorem 5.9 *An* $N \times N$ *matrix is invertible if and only if its nullspace is 0 (which is true if and only if its range is* N*-dimensional, which is true if and only if it has* N *independent rows, which is true if and only if it has* N *independent columns, which is true if and only if it has no zero eigenvalues).*

Finally, we define

Definition 5.13 *The* **rank** *of a matrix is the dimension of its range (which is the same as the maximal number of independent rows or of independent columns of the matrix).*

We can restate the previous theorem: an $N \times N$ matrix is invertible if and only if it has rank N.

5.3 Change of Basis for a Vector

Suppose we are in $\mathcal{R}^N$ or $\mathcal{C}^N$ and we want to change basis to an arbitrary (possibly non-orthogonal) basis $\mathbf{e}_i$, $i = 0, \ldots, N-1$. How do we find the coordinates of a vector in this new basis?

The answer is simple, though it's not immediately obvious why it is the right answer, but here it is: form the matrix $\mathbf{C}$ each of whose columns is one of the basis vectors $\mathbf{e}_i$. By definition of a basis, this set of vectors is independent, so $\mathbf{C}$ has N independent rows and columns and is invertible. Then the coordinates of a vector $\mathbf{v}$ in the $\mathbf{e}_i$ basis are given by $\mathbf{C}^{-1}\mathbf{v}$; that is, $\mathbf{v} = \sum_i v_i \mathbf{e}_i$ where $v_i = (\mathbf{C}^{-1}\mathbf{v})_i$.

Algebraically, this result can be derived as follows. To distinguish the coordinates from the vector $\mathbf{v}$ itself, let's use $\mathbf{b}$ to represent the vector of desired coordinates, *i.e.* we are looking for $\mathbf{b}$ such that $\mathbf{v} = \sum_i b_i \mathbf{e}_i$. Then we can rewrite $\mathbf{v} = \sum_i b_i \mathbf{e}_i$ as $\mathbf{v} = \mathbf{Cb}$ (recall Problem 3, and remember that the $\mathbf{e}_i$ are the columns of $\mathbf{C}$). This has the unique solution $\mathbf{b} = \mathbf{C}^{-1}\mathbf{v}$.

We can gain more understanding of this as follows. Let $\mathbf{f}_i^\dagger$ represent the i^{th} row of $\mathbf{C}^{-1}$. Then $\mathbf{C}^{-1}\mathbf{C} = \mathbf{1}$ translates into $\mathbf{f}_i \cdot \mathbf{e}_j = \delta_{ij}$. That is, the vector $\mathbf{f}_i$ is orthogonal to all of the basis vectors except $\mathbf{e}_i$, and it is normalized so its dot product with $\mathbf{e}_i$ is 1. So starting with $\mathbf{v} = \sum_j b_j \mathbf{e}_j$, we can take the dot product with $\mathbf{f}_i$ to find $\mathbf{f}_i \cdot \mathbf{v} = \sum_j b_j \mathbf{f}_i \cdot \mathbf{e}_j = \sum_j b_j \delta_{ij} = b_i$. So $b_i = \mathbf{f}_i \cdot \mathbf{v}$, which is just another way of stating that $(\mathbf{b})_i = (\mathbf{C}^{-1}\mathbf{v})_i$.

We can also gain a bit of geometric intuition into this result, as follows. $\mathbf{f}_i$ is orthogonal to the subspace spanned by the $N-1$ vectors $\mathbf{e}_j$, $j \neq i$; therefore, among the eigenvectors, only $\mathbf{e}_i$ has any component in the $\mathbf{f}_i$ direction. Thus the coordinate $b_i = \mathbf{f}_i \cdot \mathbf{v}$ of $\mathbf{v}$ in the $\mathbf{e}_i$ direction had better exactly account for the amount of $\mathbf{v}$ that is in the $\mathbf{f}_i$ direction, because none of the other $\mathbf{e}_j$ can contribute to this direction. Let $\hat{\mathbf{f}}_i = \mathbf{f}_i/|\mathbf{f}_i|$ be the unit vector in the $\mathbf{f}_i$ direction. The amount of $\mathbf{v}$ in the $\mathbf{f}_i$ direction is $\hat{\mathbf{f}}_i \cdot \mathbf{v}$. The amount of $b_i\mathbf{e}_i$ in the $\mathbf{f}_i$ direction is $\hat{\mathbf{f}}_i \cdot (b_i\mathbf{e}_i) = b_i\hat{\mathbf{f}}_i \cdot \mathbf{e}_i$. These two must be equal; this gives $b_i = \frac{\hat{\mathbf{f}}_i \cdot \mathbf{v}}{\hat{\mathbf{f}}_i \cdot \mathbf{e}_i} = \frac{\hat{\mathbf{f}}_i}{\hat{\mathbf{f}}_i \cdot \mathbf{e}_i} \cdot \mathbf{v} = \mathbf{f}_i \cdot \mathbf{v}$. To see the last step, note that $\frac{\hat{\mathbf{f}}_i}{\hat{\mathbf{f}}_i \cdot \mathbf{e}_i}$ points in the $\hat{\mathbf{f}}_i$ direction, and is normalized such that its dot product with $\mathbf{e}_i$ is 1; that is exactly $\mathbf{f}_i$.

So to summarize: to find the coordinate v_i in $\mathbf{v} = \sum_i v_i \mathbf{e}_i$ (we'll go back now to calling it v_i rather than b_i), find the vector $\mathbf{f}_i$ defined by (1) $\mathbf{f}_i$ is orthogonal to the subspace spanned by the $\mathbf{e}_j$ for $j \neq i$ (this defines the direction of $\mathbf{f}_i$) and (2) $\mathbf{f}_i \cdot \mathbf{e}_i = 1$ (this defines the length of $\mathbf{f}_i$). Then $v_i = \mathbf{f}_i \cdot \mathbf{v}$. This is summarized by saying that, under a change of basis to the basis $\{\mathbf{e}_i\}$, vectors are mapped $\mathbf{v} \to \mathbf{C}^{-1}\mathbf{v}$, where $\mathbf{C}$ is the matrix whose columns are $\mathbf{e}_i$.

5.4 Change of Basis for a Square Matrix: Similarity Transforms

How should a square $N \times N$ matrix transform under this change of basis? We use the same arguments as before: we'll use a prime to indicate a vector or square matrix in the new basis, *e.g.* $\mathbf{M}$, $\mathbf{v}$ are a matrix and vector in the original basis and $\mathbf{M}'$, $\mathbf{v}'$ are the same matrix and vector in the new basis. Vectors in the new basis are found by applying $\mathbf{C}^{-1}$ to vectors in the old basis. We want the action of $\mathbf{M}$ on any vector $\mathbf{v}$ to be the same in any basis, that is, $\mathbf{M}'\mathbf{v}' = (\mathbf{M}\mathbf{v})'$. This yields $\mathbf{M}'\mathbf{C}^{-1}\mathbf{v} = \mathbf{C}^{-1}(\mathbf{M}\mathbf{v})$, or $\mathbf{C}\mathbf{M}'\mathbf{C}^{-1}\mathbf{v} = \mathbf{M}\mathbf{v}$ for every vector $\mathbf{v}$. This can only be true[1] if $\mathbf{C}\mathbf{M}'\mathbf{C}^{-1} = \mathbf{M}$, or $\mathbf{M}' = \mathbf{C}^{-1}\mathbf{M}\mathbf{C}$.

This should all remind you of what we saw before. When we had an orthonormal basis, we considered the matrix $\mathbf{C}$ as the matrix all of whose columns are basis vectors (actually we called it $\mathbf{O}^\dagger$, but now we will call it $\mathbf{C}$). This was an orthogonal (or unitary) matrix, $\mathbf{C}^\dagger = \mathbf{C}^{-1}$; vectors were transformed $\mathbf{v} \to \mathbf{C}^\dagger\mathbf{v}$, and matrices were transformed $\mathbf{M} \to \mathbf{C}^\dagger\mathbf{M}\mathbf{C}$. If the bases are not orthonormal, however, then $\mathbf{C}^{-1} \neq \mathbf{C}^\dagger$; but things work much as before if we substitute $\mathbf{C}^{-1}$ for $\mathbf{C}^\dagger$. The case of an orthonormal basis was a special case of this more general rule for transforming bases.

In particular, suppose $\mathbf{M}$ has a complete basis of eigenvectors $\mathbf{e}_i$. Then a change of basis to the eigenvector basis turns $\mathbf{M}$ into a diagonal matrix $\mathbf{D}$, whose diagonal entries are just the eigenvalues of $\mathbf{M}$: $\mathbf{D} = \mathbf{C}^{-1}\mathbf{M}\mathbf{C}$ where $\mathbf{C}$ is the matrix whose columns are the eigenvectors of $\mathbf{M}$. But this means $\mathbf{M} = \mathbf{C}\mathbf{D}\mathbf{C}^{-1}$. We're going to be using this, so to be sure you don't miss it we'll say it loud:

Fact 5.1 *If* $\mathbf{M}$ *has a complete basis of eigenvectors, then* $\mathbf{M} = \mathbf{C}\mathbf{D}\mathbf{C}^{-1}$ *where* $\mathbf{D}$ *is the diago-*

[1] You may be wondering about non-square matrices. There is a problem: for a non-square $P \times N$ matrix $\mathbf{M}$, $\mathbf{v}$ and $\mathbf{M}\mathbf{v}$ live in different spaces (in $\mathcal{S}^N$ and $\mathcal{S}^P$, respectively, where $\mathcal{S}$ is the space of scalars). We've introduced the $N \times N$ matrix $\mathbf{C}^{-1}$ as a change of basis in $\mathcal{S}^N$. So $\mathbf{M}' = \mathbf{M}\mathbf{C}$ will take an N-vector from the new basis, translate it back to the old basis (by applying $\mathbf{C}$), then apply $\mathbf{M}$ to map it to $\mathcal{S}^P$. If we haven't also changed basis in $\mathcal{S}^P$, then that's the full transformation. If we've also changed basis in $\mathcal{S}^P$ by some $P \times P$ matrix $\mathbf{D}^{-1}$, then the transformed matrix would be $\mathbf{D}^{-1}\mathbf{M}\mathbf{C}$: map back to the old basis in $\mathcal{S}^N$, apply $\mathbf{M}$, then map into the new basis in $\mathcal{S}^P$.

nal matrix whose entries are the eigenvalues of $\mathbf{M}$*, and* $\mathbf{C}$ *is the matrix whose columns are the eigenvectors of* $\mathbf{M}$*.*

The transformation of matrices that we've just seen leads us to define:

Definition 5.14 *Two square matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *are said to be* **similar** *if* $\mathbf{A} = \mathbf{C}^{-1}\mathbf{B}\mathbf{C}$ *for some invertible matrix* $\mathbf{C}$*. The transform* $\mathbf{B} \to \mathbf{C}^{-1}\mathbf{B}\mathbf{C}$ *is called a* **similarity transform***.*

Many properties are preserved under similarity transforms:

Theorem 5.10 *The rank of a matrix is preserved under similarity transforms.*

Problem 5.1 *1. Using* $\det(\mathbf{A}\mathbf{B}) = (\det \mathbf{A})(\det \mathbf{B})$*, prove:*

Theorem 5.11 *The determinant of a matrix is preserved under similarity transforms.*

2. Go on to prove:

Theorem 5.12 *The eigenvalues of a matrix are preserved under similarity transforms.*

To prove this, note that the eigenvalues of $\mathbf{M}$ *are the solutions* λ *of* $\det(\mathbf{M} - \lambda\mathbf{1}) = 0$*, while those of* $\mathbf{C}^{-1}\mathbf{M}\mathbf{C}$ *are the solutions of* $\det(\mathbf{C}^{-1}\mathbf{M}\mathbf{C} - \lambda\mathbf{1}) = 0$*. You can rewrite the latter as* $\det(\mathbf{C}^{-1}(\mathbf{M} - \lambda\mathbf{1})\mathbf{C}) = 0$*. Use* $\det(\mathbf{A}\mathbf{B}) = (\det \mathbf{A})(\det \mathbf{B})$*, and the fact that* $\mathbf{C}$ *is invertible so its determinant is nonzero, to prove that this reduces to* $\det(\mathbf{M} - \lambda\mathbf{1}) = 0$*. (Note that you've actually proven something stronger: that the characteristic polynomial* $\det(\mathbf{M} - \lambda\mathbf{1})$ *is preserved under similarity transforms: all similar matrices have the same characteristic polynomial.)*

3. Prove that the eigenvectors of a matrix are preserved under similarity transforms: if $\mathbf{M}\mathbf{v} = \lambda\mathbf{v}$*, then* $(\mathbf{C}^{-1}\mathbf{M}\mathbf{C})(\mathbf{C}^{-1}\mathbf{v}) = \lambda(\mathbf{C}^{-1}\mathbf{v})$*.*

4. Prove $\mathbf{A} = \mathbf{1}$ *is preserved under similarity transforms;*

5. Prove $\mathbf{A} = 0$ *is preserved under similarity transforms;*

6. Prove "$\mathbf{A}$ *is the inverse of* $\mathbf{B}$*" is preserved under similarity transforms, i.e.* $\mathbf{A} = \mathbf{B}^{-1}$ *implies* $\mathbf{C}^{-1}\mathbf{A}\mathbf{C} = (\mathbf{C}^{-1}\mathbf{B}\mathbf{C})^{-1}$*.*

However, there is an important difference between similarity transforms and orthogonal (or unitary) transforms: similarity transforms need *not* preserve the property "$\mathbf{A}$ is the transpose (or adjoint) of $\mathbf{B}$", need *not* preserve the property "$\mathbf{A}$ is symmetric (or Hermitian)", and need *not* preserve the property "$\mathbf{A}$ is orthogonal (or unitary)".

Exercise 5.1 *Try to prove that each of these statements is preserved under a similarity transform, see where the argument breaks down, and see why it does not break down in the same place if one is considering orthogonal (or unitary) transformation. For example, if* $\mathbf{A} = \mathbf{B}^{\mathrm{T}}$*, does* $\mathbf{C}^{-1}\mathbf{A}\mathbf{C} = (\mathbf{C}^{-1}\mathbf{B}\mathbf{C})^{\mathrm{T}}$*? etc.*

These properties are preserved by orthogonal (unitary) transforms, but not by general similarity transforms. Thus, the concept that a matrix is orthogonal (unitary) or symmetric (Hermitian) only makes sense when we restrict ourselves to orthogonal (unitary) transforms. Furthermore, similarity transforms do not preserve the values of scalars, such as dot products (preserving dot products was what defined orthogonal transformations) or expressions of the form $\mathbf{x}^{\dagger}\mathbf{M}\mathbf{y}$.

5.5 Understanding General Square Matrices: Eigenvectors, Eigenvalues and Generalized Eigenspaces

When does a matrix have a complete basis of eigenvectors? To address this, we will assume we are dealing with square complex matrices – we assume the underlying scalars are the complex numbers since real matrices may have complex eigenvalues and eigenvectors.

An $N \times N$ matrix $\mathbf{M}$ always has N eigenvalues. These are, by definition, the N solutions for λ in the equation $\det(\mathbf{M} - \lambda\mathbf{1}) = 0$, which is known as the **characteristic equation** for $\mathbf{M}$. The characteristic equation is an N^{th}-order polynomial in λ and hence always has N solutions. However, these need not all be distinct: for example the second-order polynomial equation $(x-2)(x-2) = 0$ has two solutions, but both are equal to 2.

Although $\mathbf{M}$ always has N (not necessarily distinct) eigenvalues, it does not follow that $\mathbf{M}$ has N independent eigenvectors, defined as vectors $\mathbf{v}_i$ for which $(\mathbf{M} - \lambda_i\mathbf{1})\mathbf{v}_i = 0$ with λ_i a solution to $\mathbf{M}$'s characteristic equation. We have already seen one condition that guarantees N independent eigenvectors, namely if $\mathbf{M}$ is normal, $\mathbf{M}\mathbf{M}^\dagger = \mathbf{M}^\dagger\mathbf{M}$; in that case there are always N orthonormal eigenvectors. Another condition that guarantees N independent eigenvectors is if all of the eigenvalues of $\mathbf{M}$ are distinct:

Theorem 5.13 *If the $N \times N$ matrix $\mathbf{M}$ has N distinct eigenvalues, then $\mathbf{M}$ has N independent eigenvectors, one corresponding to each of the distinct eigenvalues. In this case, letting the matrix $\mathbf{C}$ be the matrix whose columns are the independent eigenvectors of $\mathbf{M}$, $\mathbf{C}^{-1}\mathbf{M}\mathbf{C}$ is a diagonal matrix $\mathbf{D}$ whose diagonal entries are the eigenvalues of $\mathbf{M}$ (and $\mathbf{M} = \mathbf{C}\mathbf{D}\mathbf{C}^{-1}$).*

Problem 5.2 *Consider the matrix $\mathbf{M} = \begin{pmatrix} 1 & 2 \\ 0 & 1+\epsilon \end{pmatrix}$ for $\epsilon \neq 0$. Show that the eigenvalues and eigenvectors are $\lambda_1 = 1$, $\mathbf{e}_1 \propto \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\lambda_2 = 1+\epsilon$, $\mathbf{e}_2 \propto \begin{pmatrix} 1 \\ \epsilon/2 \end{pmatrix}$. Let $\mathbf{C} = \begin{pmatrix} 1 & 1 \\ 0 & \epsilon/2 \end{pmatrix}$. The inverse of a 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with determinant $D = ad - bc$ is given by $\begin{pmatrix} d/D & -b/D \\ -c/D & a/D \end{pmatrix}$. Use this to verify that $\mathbf{C}^{-1}\mathbf{M}\mathbf{C} = \begin{pmatrix} 1 & 0 \\ 0 & 1+\epsilon \end{pmatrix}$.*

Problem 5.3 *Recall from Problem 3 that if $\mathbf{M}$ has columns $\mathbf{c}_i$ and $\mathbf{N}$ has rows $\mathbf{r}_i^\dagger$ then $\mathbf{M}\mathbf{N} = \sum_i \mathbf{c}_i\mathbf{r}_i^\dagger$. Consider the equation $\mathbf{M} = \mathbf{C}\mathbf{D}\mathbf{C}^{-1}$. The columns of $\mathbf{C}$ are the eigenvectors $\mathbf{e}_i$ of $\mathbf{M}$. Let $\mathbf{f}_i^\dagger$ be the rows of $\mathbf{C}^{-1}$; we saw in section 5.3 that the $\mathbf{f}_i$ are defined by $\mathbf{f}_i^\dagger\mathbf{e}_j = \delta_{ij}$, and that $\mathbf{f}_i^\dagger\mathbf{v}$ gives the coordinate of $\mathbf{v}$ along the i^{th} eigenvector $\mathbf{e}_i$. $\mathbf{D}$ is a diagonal matrix with entries λ_i, the eigenvalues of $\mathbf{M}$.*

Now put this all together and show that $\mathbf{M} = \mathbf{C}\mathbf{D}\mathbf{C}^{-1}$ can be rewritten $\mathbf{M} = \sum_i \lambda_i\mathbf{e}_i\mathbf{f}_i^\dagger$ (hint: all you have to add is to show that the rows of $\mathbf{D}\mathbf{C}^{-1}$ are $\lambda_i\mathbf{f}_i^\dagger$, or that the columns of $\mathbf{C}\mathbf{D}$ are $\lambda_i\mathbf{e}_i$). In words: when $\sum_i \lambda_i\mathbf{e}_i\mathbf{f}_i^\dagger$ is applied to a vector $\mathbf{v}$, the $\mathbf{f}_i^\dagger$ finds the coordinate of $\mathbf{v}$ along the $\mathbf{e}_i$ direction; this is then multiplied by $\mathbf{e}_i$, and scaled by λ_i. Repeat this for each of the eigenvectors (sum over i) and you obtain the action of $\mathbf{M}$ on $\mathbf{v}$.

Exercise 5.2 *From the formulation $\mathbf{M} = \sum_j \lambda_j\mathbf{e}_j\mathbf{f}_j^\dagger$ we can draw a further conclusion: the $f_i^\dagger$ are the* left eigenvectors *of $\mathbf{M}$, that is, $\mathbf{f}_i^\dagger\mathbf{M} = \lambda_i\mathbf{f}_i^\dagger$, while the $\mathbf{e}_i$ are the right eigenvectors of $\mathbf{M}$, $\mathbf{M}\mathbf{e}_i = \lambda_i\mathbf{e}_i$. Show this by computing $\mathbf{f}_i^\dagger\mathbf{M}$ and $\mathbf{M}\mathbf{e}_i$ with $\mathbf{M}$ written as $\sum_j \lambda_j\mathbf{e}_j\mathbf{f}_j^\dagger$, recalling that $\mathbf{f}_j^\dagger\mathbf{e}_i = \delta_{ji}$. Thus, $\mathbf{M}$ has a set of left eigenvectors $\mathbf{f}_i^\dagger$ and of right eigenvectors $\mathbf{e}_i$; neither set is*

orthonormal, but the two are "mutually orthonormal", meaning $\mathbf{f}_j^\dagger \mathbf{e}_i = \delta_{ji}$. *For the special case of normal matrices, the* $\mathbf{e}_i$ *are orthonormal and so* $\mathbf{f}_i^\dagger = \mathbf{e}_i^\dagger$ *– the left and right eigenvectors are identical.*

Old linear algebra books talked about left and right eigenvectors, but it seems to have gone out of fashion. Nowadays they just talk about eigenvectors, meaning right eigenvectors; they form $\mathbf{C}$ *as the matrix whose columns are these eigenvectors, and use* $\mathbf{C}^{-1}$ *as above without noting that its rows are the left eigenvectors. That probably makes sense, as there's no particular use of the left eigenvectors except to form* $\mathbf{C}^{-1}$*, but it still seems an interesting point to notice. But outside of this exercise we will follow the crowd and just say "eigenvectors", meaning "right eigenvectors".*

If two or more eigenvalues of $\mathbf{M}$ are equal to one another, we say these eigenvalues are **degenerate**; if k eigenvalues have the same value, we say they have a k-degeneracy. In this case $\mathbf{M}$ may be "missing" some eigenvectors – corresponding to k degenerate eigenvalues, there may be any number from 1 to k independent eigenvectors. (However if $\mathbf{M}$ is normal, then we are guaranteed there are no missing eigenvectors.) In this case we can define the "generalized eigenspace" as the set of solutions $\mathbf{v}$ to the equation $(\mathbf{M}-\lambda_d\mathbf{1})^k\mathbf{v}=0$, where λ_d is the k-degenerate eigenvalue. There will always be k independent elements of the generalized eigenspace of λ_d, and they will be independent of the (generalized) eigenvectors corresponding to other eigenvalues. Furthermore there is always at least one eigenvector in the generalized eigenspace: if $\mathbf{v}$ is in the generalized eigenspace, $(\mathbf{M}-\lambda_d\mathbf{1})^k\mathbf{v}=0$, which we can rewrite $(\mathbf{M}-\lambda_d\mathbf{1})\left[(\mathbf{M}-\lambda_d\mathbf{1})^{k-1}\mathbf{v}\right]=0$, so $(\mathbf{M}-\lambda_d\mathbf{1})^{k-1}\mathbf{v}$ is either an eigenvector of $\mathbf{M}$ with eigenvalue λ_d, or else it is zero (in which case $(\mathbf{M}-\lambda_d\mathbf{1})^{k-2}\mathbf{v}$ is either an eigenvector, or zero, and so on, remembering that when we get down to $(\mathbf{M}-\lambda_d\mathbf{1})\mathbf{v}=0$, it means $\mathbf{v}$ is an eigenvector).

Problem 5.4 *Consider the matrix* $\mathbf{M}=\begin{pmatrix}1&2\\0&1\end{pmatrix}$. *The characteristic equation for* $\mathbf{M}$ *is* $(\lambda-1)^2=0$*, which has two roots that are both 1. One eigenvector is* $\begin{pmatrix}1\\0\end{pmatrix}$*, but there is no second eigenvector (confirm this). The generalized eigenspace is given by the solutions* $\mathbf{v}$ *to* $(\mathbf{M}-\mathbf{1})^2\mathbf{v}=0$. *But* $(\mathbf{M}-\mathbf{1})^2=0$ *(confirm this) so every vector* $\mathbf{v}$ *satisfies this equation; so in particular we can take the second element of the generalized eigenspace to be* $\begin{pmatrix}0\\1\end{pmatrix}$. *Note that the same arguments would apply to any matrix* $\mathbf{M}=\begin{pmatrix}1&a\\0&1\end{pmatrix}$ *with* $a\neq 0$.

(Note that if we break the degeneracy between the eigenvalues, then there are two independent eigenvectors. In Problem 5.2 we considered the matrix $\mathbf{M}=\begin{pmatrix}1&2\\0&1+\epsilon\end{pmatrix}$*; those results showed that, so long as* $\epsilon\neq 0$*, the eigenvalue degeneracy is broken, and there are two distinct and independent eigenvectors. For* $\epsilon=0$*, however, the two eigenvalues become identical, as do the two eigenvectors.)*

As an aside – this may be too compressed to really make sense to you, in which case don't get stuck on it: One can gain a bit of intuition for why some matrices may have missing eigenvectors as follows. The **Cayley-Hamilton** theorem says that each matrix satisfies its own characteristic equation. That is, suppose the characteristic equation of $\mathbf{M}$, $\det(\mathbf{M}-\lambda\mathbf{1})=0$, gives the N-th order polynomial equation $\sum_{i=0}^N a_i\lambda^i=0$, where necessarily $a_N=1$; $\sum_{i=0}^N a_i\lambda^i$ is then called the **characteristic polynomial** of $\mathbf{M}$. Then the Cayley-Hamilton theorem states that $\sum_{i=0}^N a_i\mathbf{M}^i=0$. If the N eigenvalues of $\mathbf{M}$ are λ_i, $i=0,\ldots,N-1$, then the characteristic polynomial can

be written as $(\lambda - \lambda_0)(\lambda - \lambda_1)\dots(\lambda - \lambda_{N-1})$. So the Cayley-Hamilton theorem tells us that $(\mathbf{M}-\lambda_0)(\mathbf{M}-\lambda_1)\dots(\mathbf{M}-\lambda_{N-1}) = 0$. $(\mathbf{M}-\lambda_0)$ will give zero when applied to an eigenvector with eigenvalue λ_0; $(\mathbf{M}-\lambda_1)$ will give zero when applied to an eigenvector with eigenvalue λ_1, etc. So in essence, what the Cayley-Hamilton theorem suggests (and what in fact can be proved) is that the whole vector space can be decomposed into a sum of the generalized eigenspace associated with λ_0, the generalized eigenspace associated with λ_1, ... – every vector is a linear combination of elements of these generalized eigenspaces, and so $(\mathbf{M} - \lambda_0)(\mathbf{M} - \lambda_1)\dots(\mathbf{M} - \lambda_{N-1})$ applied to any vector gives zero. If an eigenvalue λ_i is not degenerate, it has a one-dimensional eigenspace, and elements $\mathbf{e}_i$ of that eigenspace satisfy $(\mathbf{M} - \lambda_i)\mathbf{e}_i = 0$ – they are eigenvectors of $\mathbf{M}$. But if an eigenvalue λ_j is k-degenerate, and $\mathbf{e}_j$ is an element of the associated generalized eigenspace, then all that we know from the Cayley-Hamilton theorem is that $(\mathbf{M}-\lambda_j)^k\mathbf{e}_j = 0$. There is no guarantee that these elements are eigenvectors, only that they form part of the generalized eigenspace of λ_j.

5.6 Linear Differential Equations With General Square Matrices

We can now write down the general solution to first-order linear differential equations:

Theorem 5.14 *Consider the equation $\frac{d}{dt}\mathbf{v} = \mathbf{M}\mathbf{v}$. Let λ_i be the eigenvalues of $\mathbf{M}$, with corresponding eigenvectors, or if need be generalized eigenvectors, $\mathbf{e}_i$. Let $\mathbf{C}$ be the matrix whose columns are the $\mathbf{e}_i$. Let the initial condition be $\mathbf{v}(0)$, and write $\mathbf{v}(0) = \sum_i v_i(0)\mathbf{e}_i$ with $v_i(0) = \left(\mathbf{C}^{-1}\mathbf{v}(0)\right)_i$. Then if $\mathbf{M}$ has a complete basis of eigenvectors, the solution is*

$$\mathbf{v}(t) = \sum_i \mathbf{e}_i v_i(0) e^{\lambda_i t} \tag{5.1}$$

If $\mathbf{M}$ has a k-degenerate subspace with eigenvalue λ_d that is missing some eigenvectors, all factors of $e^{\lambda_d t}$ should be replaced by $e^{\lambda_d t} + \sum_{p=1}^{k-1} c_p t^p e^{\lambda_d t}$ where the c_p are constants to be determined.

When no eigenvectors are missing, this is exactly the formula we had been using previously – the only difference is now the eigenvectors are not necessarily orthonormal. The origin of the mysterious factors of t^p in the case of missing eigenvectors can be understood by using a somewhat more powerful approach, which also provides a more elegant and powerful way to find the solutions in the case of missing eigenvectors – look up the Jordan normal form and the solution in terms of the exponential of a matrix in a linear algebra book. But for our purposes this characterization is sufficient – missing eigenvectors rarely if ever come up in real life.

If there are missing eigenvectors, one can simply write down the terms given by theorem 5.14 and solve for their coefficients:

Problem 5.5 *Consider again the matrix $\mathbf{M} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, and consider the equation $\frac{d}{dt}\mathbf{v} = \mathbf{M}\mathbf{v}$. Recall the eigenvalues are both 1, the one eigenvector is $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, and the other generalized eigenvector can be taken to be $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Write $\mathbf{v}(t) = \begin{pmatrix} v_1(0)e^t + k_1 te^t \\ v_2(0)e^t + k_2 te^t \end{pmatrix}$ and solve for the constants k_1 and k_2. You should arrive at the solution $\mathbf{v}(t) = \begin{pmatrix} v_1(0)e^t + 2v_2(0)te^t \\ v_2(0)e^t \end{pmatrix}$.*

Similarly, if $\mathbf{M}$ has a complete basis of eigenvectors, the solution to the inhomogeneous equation $\frac{d}{dt}\mathbf{v} = \mathbf{M}\mathbf{v} + \mathbf{h}(t)$ can be written just as before:

$$\mathbf{v}(t) = \sum_i \mathbf{e}_i e^{\lambda_i t}\left[\int_0^t ds\, e^{-\lambda_i s} h_i(s) + v_i(0)\right] \tag{5.2}$$

where $v_i(0) = \left(\mathbf{C}^{-1}\mathbf{v}(0)\right)_i$, $h_i(s) = \left(\mathbf{C}^{-1}\mathbf{h}(s)\right)_i$, and again $\mathbf{C}$ is the matrix whose columns are the eigenvectors of $\mathbf{M}$. If $\mathbf{M}$ has a k-degenerate subspace with eigenvalue λ_d that is missing some eigenvectors, $e^{\lambda_d t}$ should be replaced by a linear combination of $t^p e^{\lambda_d t}$ for $p = 0, \ldots, k-1$, and $e^{-\lambda_d s}$ should be replaced by a linear combination of $s^p e^{-\lambda_d s}$ for $p = 0, \ldots, k-1$.

5.7 Non-Square Matrices: The Singular Value Decomposition

We now turn to the singular value decomposition (SVD), a powerful method of decomposing a matrix that can be applied to any matrix, square or rectangular. The SVD defines a set of *singular values* that are associated with the matrix. For a (square) normal matrix, the singular values σ_i are simply the absolute values of the eigenvalues λ_i of the matrix, $\sigma_i = \sqrt{\lambda_i^* \lambda_i}$. In particular, for a matrix that is hermitian, so that all of its eigenvalues are real and its eigenvectors are orthonormal, we shall see that the SVD is essentially identical to the eigenvector/eigenvalue decomposition; but for non-hermitian square matrices the SVD and the eigenvector/eigenvalue decomposition are different (and for non-square matrices, the concept of eigenvector is not even defined).

Theorem 5.15 Singular value decomposition (SVD): *Every $P \times N$ matrix $\mathbf{M}$ can be decomposed as*

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\dagger \tag{5.3}$$

where $\mathbf{U}$ is a $P \times P$ unitary matrix, $\mathbf{V}$ is an $N \times N$ unitary matrix, and $\mathbf{\Sigma}$ is a diagonal $P \times N$ matrix (meaning that it is nonzero only along the diagonal, which has length $\min\{P, N\}$) whose diagonal entries are real, non-negative, and are known as the **singular values** *of $\mathbf{M}$. If the rank of $\mathbf{M}$ is r, there are r nonzero singular values.*

We assume the matrices are arranged so that the singular values are ordered from largest to smallest. Then let $\mathbf{v}_i$, $i = 0, \ldots, N-1$ be the columns of $\mathbf{V}$; $\mathbf{u}_i$, $i = 0, \ldots, P-1$ be the columns of $\mathbf{U}$; and σ_i, $i = 0, \ldots, r-1$ be the nonzero singular values. Then we can also write the SVD as

$$\mathbf{M} = \sum_{i=0}^{r-1} \sigma_i \mathbf{u}_i \mathbf{v}_i^\dagger \tag{5.4}$$

The formulation of Eq. 5.4 should make clear that the first r columns of $\mathbf{V}$ form a basis for the space spanned by the rows of $\mathbf{M}$; the first r columns of $\mathbf{U}$ form a basis for the range of $\mathbf{M}$; and $\mathbf{M}$ maps one basis into the other, $\mathbf{M}\mathbf{v}_i = \sigma_i \mathbf{u}_i$. Thus, the SVD finds two orthonormal bases – one for the rowspace of $\mathbf{M}$, one for the range – such that one basis is mapped into the other by $\mathbf{M}$; and the singular values tell how much each basis vector is stretched or shrunken under that mapping.

The SVD should remind you of the eigenvector decomposition for a Hermitian matrix: in that case $\mathbf{M} = \mathbf{U}\mathbf{D}\mathbf{U}^\dagger = \sum_i \lambda_i e_i e_i^\dagger$, where $\mathbf{D}$ is the diagonal matrix of the real eigenvalues λ_i and $\mathbf{U}$ is the unitary matrix whose columns are the eigenvectors $\mathbf{e}_i$. For a Hermitian matrix, the SVD *is* the eigenvector decomposition; for negative eigenvalues $\lambda_i < 0$, one defines $\sigma_i = -\lambda_i$, $\mathbf{v}_i = \mathbf{e}_i$, $\mathbf{u}_i = -\mathbf{e}_i$, but otherwise the decompositions are identical. Geometrically, the eigenvector decomposition $\mathbf{M} = \mathbf{U}\mathbf{D}\mathbf{U}^\dagger$ means rotate the $\mathbf{e}_i$ basis vectors to be the coordinate axes; scale these coordinates ($\mathbf{D}$); and then do the inverse rotation, taking the coordinate axes back to the $\mathbf{e}_i$. $\mathbf{M}$ thus maps each $\mathbf{e}_i$ to a scaled version of itself. The SVD decomposition $\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\dagger$ means rotate the $\mathbf{v}_i$ basis vectors to be the coordinate axes; scale these coordinates, and possibly embed the nonzero coordinates in a space of a different dimension; and then rotate the coordinate axes back to a different basis, the $\mathbf{u}_i$ (Fig. 5.1). $\mathbf{M}$ thus maps each $\mathbf{v}_i$ to a scaled version of the corresponding $\mathbf{u}_i$. In the case of a Hermitian matrix, the two bases are identical except possibly for signs, but more generally this is not the case.

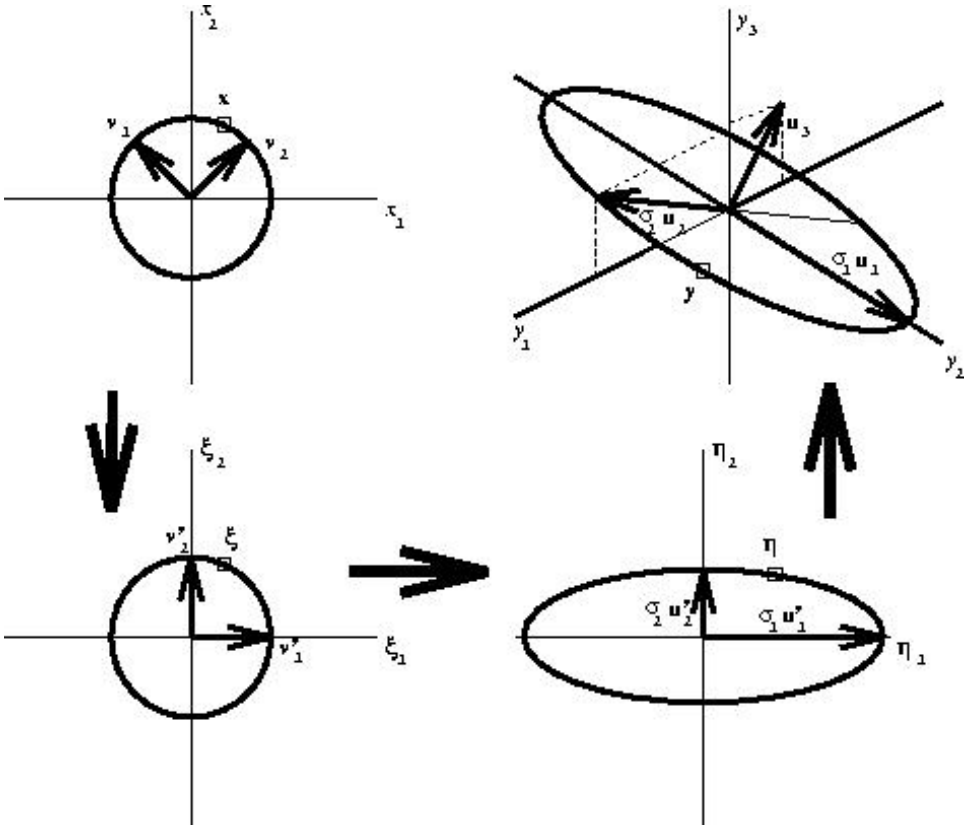

Figure 5.1: **The Singular Value Decomposition**
The singular value decomposition can be understood as the composition of three actions: rotating the vectors $\mathbf{v}_i$ to the coordinate axes; scaling these axes by the singular values; and then rotating these axes to coincide with the vectors $\mathbf{u}_i$ in the output space. The result of this is to map vectors on the unit circle in the rowspace (the space spanned by the rows; the portion of the domain that is orthogonal to the nullspace) into the principal axes of a hyperellipse in the range. This figure was stolen, with thanks, from http://www.stanford.edu/class/cs205/notes/book/node18.html.

This geometric interpretation of the SVD is illustrated in Fig. 5.1. This also shows another aspect of the SVD. For any matrix $\mathbf{M}$, the unit sphere (the sphere of radius 1) in the space spanned by the rows of $\mathbf{M}$ is mapped to a hyperellipse (a set of vectors $\mathbf{x}$ defined by $\mathbf{x}^\dagger\mathbf{Q}\mathbf{x} = 1$ for some Hermitian matrix $\mathbf{Q}$; in this case $\mathbf{Q} = \sum_i \frac{\mathbf{u}_i\mathbf{u}_i^\dagger}{\sigma_i^2}$) in the range of $\mathbf{M}$. The $\mathbf{u}_i$ are the major axis, first minor axis, second minor axis, etc. of the hyperellipse, and the $\mathbf{v}_i$ are the unit vectors in the domain of $\mathbf{M}$ that map to these hyperellipse axes in the range of $\mathbf{M}$. The σ_i are the radii of the axes of the hyperellipse.

The SVD is often used to break down a function of two variables into a sum of products of functions of one variable. For example, the response of a visual cortical neuron may depend on stimulus orientation θ and on time t (the time between stimulus and response), and one wishes to know if the orientation tuning changes as a function of time. If the response can be written as a function of orientation times a function of time, then the orientation dependence is the same at all times. Thus one takes the response function $R(\theta, t)$, discretizes it as a matrix R_{ij} representing the response to orientation θ_i at time t_j, and does an SVD. Each term in $\mathbf{R} = \sum_{i=0}^{r-1} \sigma_i \mathbf{u}_i \mathbf{v}_i^\dagger$ is a product of a function of orientation (the $\mathbf{u}_i$) times a function of time (the $\mathbf{v}_i^\dagger$) (this can be seen from $R_{ij} = \sum_{k=0}^{r-1} \sigma_k (\mathbf{u}_k)_i (\mathbf{v}_k^*)_j$ – the i dependence is in the $\mathbf{u}$, the j dependence is in the $\mathbf{v}$). If the first singular value is much larger than all the others, then $\mathbf{R}$ is reasonably described by the first term alone, which means it is reasonably described as the product of a function of orientation times a function of time. In genomics, one may have a response in each of 10,000 genes across 50 arrays, each representing a different experimental condition; SVD expresses the responses as a sum of products of a function of the gene times a function of the array.

To perform the SVD, one deals with the matrices $\mathbf{M}\mathbf{M}^\dagger$ and $\mathbf{M}^\dagger\mathbf{M}$. $\mathbf{M}\mathbf{M}^\dagger$ is a $P \times P$ hermitian matrix: $(\mathbf{M}\mathbf{M}^\dagger)^\dagger = \mathbf{M}\mathbf{M}^\dagger$ (recall that $(\mathbf{A}\mathbf{B})^\dagger = \mathbf{B}^\dagger\mathbf{A}^\dagger$). Similarly, $\mathbf{M}^\dagger\mathbf{M}$ is an $N \times N$ hermitian matrix. Thus, each matrix has real eigenvalues and has a complete orthonormal basis of eigenvectors. Furthermore the eigenvalues of each are positive or zero: if $\mathbf{e}_i$ is an eigenvector of $\mathbf{M}^\dagger\mathbf{M}$ with eigenvalue λ_i, then $\mathbf{e}_i^\dagger\mathbf{M}^\dagger\mathbf{M}\mathbf{e}_i = \lambda_i$, but also $\mathbf{e}_i^\dagger\mathbf{M}^\dagger\mathbf{M}\mathbf{e}_i = (\mathbf{M}\mathbf{e}_i)^\dagger\mathbf{M}\mathbf{e}_i = |\mathbf{M}\mathbf{e}_i|^2 \geq 0$, so $\lambda_i \geq 0$; a similar argument applies to $\mathbf{M}\mathbf{M}^\dagger$.

Because it is hermitian, $\mathbf{M}\mathbf{M}^\dagger = \mathbf{S}\mathbf{D}\mathbf{S}^\dagger$ for some diagonal matrix $\mathbf{D}$, which contains the real eigenvalues of $\mathbf{M}\mathbf{M}^\dagger$, and some unitary matrix $\mathbf{S}$, whose columns contain the eigenvectors of $\mathbf{M}\mathbf{M}^\dagger$. But if we compute $\mathbf{M}\mathbf{M}^\dagger$, we find $\mathbf{M}\mathbf{M}^\dagger = (\mathbf{U}\mathbf{\Sigma}\mathbf{V}^\dagger)(\mathbf{V}\mathbf{\Sigma}^\dagger\mathbf{U}^\dagger) = \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^\dagger\mathbf{U}^\dagger$. $\mathbf{\Sigma}\mathbf{\Sigma}^\dagger$ is a $P \times P$ diagonal matrix, so this is precisely the eigenvector/eigenvalue decomposition of $\mathbf{M}\mathbf{M}^\dagger$. Thus, the columns of $\mathbf{U}$ are the eigenvectors of $\mathbf{M}\mathbf{M}^\dagger$, and the nonzero singular values are the positive squareroots of the nonzero eigenvalues of $\mathbf{M}\mathbf{M}^\dagger$. Similarly, by writing $\mathbf{M}^\dagger\mathbf{M} = \mathbf{V}\mathbf{\Sigma}^\dagger\mathbf{\Sigma}\mathbf{V}^\dagger$, we see that the columns of $\mathbf{V}$ are the eigenvectors of $\mathbf{M}^\dagger\mathbf{M}$, and the nonzero singular values are the positive squareroots of the nonzero eigenvalues of $\mathbf{M}^\dagger\mathbf{M}$. Thus, to perform the SVD, one does an eigenvector decomposition of $\mathbf{M}\mathbf{M}^\dagger$ and of $\mathbf{M}^\dagger\mathbf{M}$.

Note that if $\mathbf{M}$ is a unitary matrix, then $\mathbf{M}\mathbf{M}^\dagger = \mathbf{M}^\dagger\mathbf{M} = \mathbf{1}$. In this case, the SVD is not of much use, since any vector is an eigenvector of $\mathbf{1}$, so the choice of $\mathbf{U}$ and $\mathbf{V}$ are largely unconstrained. Furthermore the eigenvalues of a unitary matrix have absolute value 1, so $\mathbf{\Sigma} = \mathbf{1}$ for a unitary matrix. Thus, one form of the SVD of a unitary matrix is $\mathbf{M} = \mathbf{M}\mathbf{1}\mathbf{1}$; so for unitary matrices, the SVD is not useful.

Problem 5.6 *1. Derive the SVD of the* 2×1 *matrix* $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

2. Derive the SVD of the matrix $\begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix}$.

Exercise 5.3 *Suppose* $\mathbf{M}^\dagger\mathbf{M}\mathbf{x} = 0$. *Then* $\mathbf{x}^\dagger\mathbf{M}^\dagger\mathbf{M}\mathbf{x} = 0$*; but* $\mathbf{x}^\dagger\mathbf{M}^\dagger\mathbf{M}\mathbf{x} = (\mathbf{M}\mathbf{x})^\dagger(\mathbf{M}\mathbf{x}) = |\mathbf{M}\mathbf{x}|^2$. *Conclude that if* $\mathbf{M}^\dagger\mathbf{M}\mathbf{x} = 0$*, then* $\mathbf{M}\mathbf{x} = 0$*; the converse is obviously true. So* $\mathbf{M}^\dagger\mathbf{M}$ *has the same nullspace as* $\mathbf{M}$*. The rowspace (the space spanned by the rows) is the subspace orthogonal to the nullspace; so if two matrices have the same nullspace, they have the same rowspace. Thus the eigenvectors of* $\mathbf{M}^\dagger\mathbf{M}$ *with nonzero eigenvalue form a basis for the rowspace of* $\mathbf{M}^\dagger\mathbf{M}$ *and thus also form a basis for the rowspace of* $\mathbf{M}$*.*

Similarly, show that if $\mathbf{x}\mathbf{M}\mathbf{M}^\dagger = 0$*, then* $\mathbf{x}\mathbf{M} = 0$*. Conclude that* $\mathbf{M}\mathbf{M}^\dagger$ *and* $\mathbf{M}$ *have the same column space (the space spanned by the columns), so that the eigenvectors of* $\mathbf{M}\mathbf{M}^\dagger$ *with nonzero eigenvalue form a basis for the range of* $\mathbf{M}$*.*