Math 520 Learning Objectives
Spring 2017
Brown University
Samuel S. Watson

Week 1 (23 Jan through 27 Jan)

- Definition of a system of linear equations, definition of a solution of a linear system, elementary row operations (switch rows, scale a row, and add to one row a multiple of another)
- Augmented matrix and matrix of coefficients of a system of equations, row echelon form (all entries below a leading entry are zero, and rows are sorted by the column of their leading entry), row reduction algorithm (move a row with leftmost leading entry to the top, and use that entry to zero-out the entries below it; repeat with remaining rows)
- Solving a system by back-substitution, checking consistency of a system (no rows of the form
$[0 \quad 0 \quad \cdots \quad \blacksquare]$)

Week 2 (30 Jan through 3 Feb)

- Reduced row echelon form (row echelon plus leading entries all 1 plus all zeros above each leading entry),
- Solution of a system in terms of free variables
- Vector addition and scalar multiplication
- Definition of a linear combination, definition of the span of a collection of vectors
- Matrix form $A\mathbf{x} = \mathbf{b}$ of a system of linear equations, properties of the matrix-vector product
- The columns of an $m \times n$ matrix span $\mathbb{R}^m$ if and only if the matrix has a pivot entry in every row

Week 3 (6 Feb through 10 Feb)

- Basic facts about homogeneous and nonhomogeneous systems: every homogeneous system $A\mathbf{x} = \mathbf{b}$ has at least the trivial solution $\mathbf{x} = \mathbf{0}$, and nontrivial solutions as well if A has some non-pivot column
- A nonhomogeneous system can be inconsistent, but if it's consistent then its solution takes the form (particular solution) + (solution of homogeneous system)
- Applications: balancing economies, balancing chemical reactions, and traffic networks
- Linear independence: a list of vectors is linearly dependent if $\mathbf{0}$ is a nontrivial linear combination of those vectors; alternatively, if one of the vectors is in the span of the others. Geometrically, this means that the vectors are collinear (for two vectors) or coplanar (for three vectors)

Week 4 (13 Feb through 17 Feb)

- Linear transformations map equally spaced lines to equally spaces lines (or points) and the origin to the origin. Basic examples in 2D: rotating, scaling, shearing, projecting, and reflecting.
- Multiplying a vector $\mathbf{x} \in \mathbb{R}^n$ on the left by a matrix $A \in \mathbb{R}^{m \times n}$ represents a linear transformation, and every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be represented that way
- The matrix representing a given linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^m$ is obtained by letting its columns be the vectors to which $\mathbf{e}_1 = (1,0)$ and $\mathbf{e}_2 = (0,1)$ are mapped (and similarly for any $\mathbb{R}^n$, e.g., for $\mathbb{R}^3$ we use $\mathbf{e}_1 = (1,0,0)$, $\mathbf{e}_2 = (0,1,0)$, and $\mathbf{e}_3 = (0,0,1)$)
- A matrix transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is surjective (i.e., every vector in $\mathbb{R}^m$ has a vector in $\mathbb{R}^n$ mapping to it) iff the columns of the matrix span $\mathbb{R}^m$, and it's injective (i.e., no two vectors in $\mathbb{R}^n$ map to the same output) iff all the columns of the matrix are pivot columns
- A linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ can only be bijective if $m = n$
- Matrix multiplication is defined for an $m \times n$ matrix times an $n \times p$ matrix; the result is an $m \times p$ matrix which represents the *composition* of the matrices' corresponding linear transformations: the (i,j)th entry of the product is the ith row of the first matrix dotted with the jth column of the second matrix (where 'dot' means multiply in pairs and sum). Matrix multiplication is not commutative
- This table summarizes some helpful relationships; A is an $m \times n$ matrix

A has a pivot in every row	$A\mathbf{x} = \mathbf{b}$ has at least one solution, for all $\mathbf{b} \in \mathbb{R}^m$	$T(\mathbf{x}) = A\mathbf{x}$ is surjective (onto)
A has a pivot in every column	$A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{x} = \mathbf{0}$	$T(\mathbf{x}) = A\mathbf{x}$ is injective (one-to-one)

Week 5 (20 Feb through 24 Feb)

- The inverse of a bijective linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$ is again a linear transformation
- To find the inverse of A, we find the vectors that A maps to $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_n$
- This is the same as row-reducing the augmented matrix $[A|I]$, where I is the $n \times n$ identity matrix, and reading off the resulting matrix right of the bar. If the matrix left of the bar isn't the identity, then A wasn't invertible
- $(AB)^{-1} = B^{-1}A^{-1}$
- 2×2 matrix inversion formula: $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ (not invertible if $ad - bc = 0$)
- You can solve equations whose variables represent matrices like you would if they represent numbers, except: (a) never use division: use matrix inverses instead, and (b) watch out for non-commutativity. So like $C + BA^{-1}X = 0$ implies $X = -AB^{-1}C$

Week 6 (27 Feb through 3 Mar)

- Definition of a linear subspace of $\mathbb{R}^n$ (contains $\mathbf{0}$, closed under scalar multiplication and vector addition)
- The span of any list of vectors is a linear subspace
- The *column space* of a matrix is the span of its columns
- The *null space* of a matrix is the solution set of its homogeneous equation
- A *basis* of a linear subspace is a linearly indepent spanning list
- To find a basis for Nul A, solve the homogeneous system and write the solution in span form
- To find a basis for Col A, select all the pivot columns from A (not from the rref of A!)
- Given a basis of a subspace, every vector in that subspace can be written uniquely as a linear combination of those basis vectors
- Every basis of a given linear subspace has the same number of vectors in it, and this number is called the *dimension* of the subspace
- The *rank* of a matrix is the dimension of the column space, and its *nullity* is the dimension of its null space
- The rank-nullity theorem tells us that rank plus nullity equals number of columns

Week 7 (6 Mar through 10 Mar)

- A linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^n$ scales all n-dimensional volumes by the same factor: the (absolute value of the) *determinant* of T
- The sign of the determinant tells us whether the linear transformation reverses orientations, meaning that a small counterclockwise loop drawn on the outside of a closed surface maps to a counterclockwise loop (+) or a clockwise loop (-)
- The determinant of a 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $ad - bc$
- The determinant of a linear transformation can be computed from its matrix using cofactor expansions: (1) choose a row or column, (2) for each entry in that row or column, multiply (i) a checkerboard-alternating sign, (ii) the matrix entry, and (iii) the determinant of the matrix obtained by removing that entry's row and column from the matrix. Finally, (3) sum all the terms from step 2
- A rook arrangement is a selection of n positions in a $n \times n$ matrix with the property that no two selected positions are in the same row or column
- The sign of a rook arrangement is the number of column switches it takes to get there from the main-diagonal rook arrangement
- The determinant of a matrix can also be computed as a sum over all rook arrangements on the matrix of the product of the entries in the rook arrangement's positions times the sign of the rook arrangement
- Row switching multiplies the determinant of a matrix by -1, row scaling scales the determinant, and row-add operations don't alter the determinant
- The determinant of an upper triangular matrix (all zeros below the main diagonal) is the product of the entries on the main diagonal (why? the main-diagonal rook arrangement is the only nonzero term in the rook expansion of the determinant)
- The computationally efficient way to compute determinants is to row reduce to echelon form, multiply down the diagonal, and account for any changes you made to the determinant whilst row reducing
- $\det AB = \det A \det B$ if A and B are both $n \times n$, because scaling volumes by a factor of k and then by a factor of ℓ is equivalent to scaling volumes by a factor of $k\ell$
- $\det A^{-1} = (\det A)^{-1}$, because if A scales volumes by a factor of k, then the inverse of A must scale volumes by a factor of $1/k$
- Cramer's rule says that to find the ith component of the solution of $A\mathbf{x} = \mathbf{b}$, we can replace the ith column of A with $\mathbf{b}$, find the determinant of that matrix, and divide the result by $\det A$

Week 8 (13 Mar through 17 Mar)

- A vector space is a set V together with an operation $+$ which tells us how to add elements of V and an operation which tells us how to multiply elements of V by real numbers. These operations are required to satisfy a list of 10 basic properties, like commutativity and associativity of addition, etc.
- Working with abstract vector spaces is useful because it allows us to develop concepts we learned for $\mathbb{R}^n$ in a way that permits their application to a variety of mathematical objects whose elements are not n-tuples of real numbers
- $\mathbb{P}_n$ is the space of polynomials of degree at most n, with usual polynomial addition and scalar multiplication
- $C([0,1])$ is the space of continuous functions from $[0,1]$ to $\mathbb{R}$, with usual function addition and scalar multiplication
- A subspace of a vector space V is a subset of V which contains the zero vector and is closed under the vector operations of addition and scalar multiplication
- The span of a list of vectors in V is always a subspace of V
- A linear transformation T from V to W is a function which satisfies $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ and $T(c\mathbf{u}) = cT(\mathbf{u})$ for all real numbers c and all $\mathbf{u}, \mathbf{v} \in T$
- The *kernel* of a linear transformation from V to W is the set of vectors that map to the zero vector in W
- The *range* of a linear transformation from V to W is the set of all vectors in W that get mapped to
- Linear independence of a list of vectors in an abstract vector space has the same definition as in $\mathbb{R}^n$
- The linear dependence lemma: a list of vectors is linearly dependent if and only if there is some vector in the list which is in the span of the vectors appearing *before* it in the list. Removing this vector does not change the span of the list
- Any linearly independent list in a vector space V is no longer than any list of vectors that spans V. Abbreviated: **linearly independent lists are no longer than spanning lists**

Week 9 (20 Mar through 24 Mar)

- Each vector in a vector space V can be uniquely represented as a linear combination of the vectors in a given basis of V
- The weights in that linear combination are called the *coordinates* of the vector with respect to the given basis
- The coordinates of $\mathbf{x}$ with respect to $\mathcal{B}$ are denoted $[\mathbf{x}]_{\mathcal{B}}$
- To convert the standard coordinates of $\mathbf{v} \in \mathbb{R}^n$ to coordinates with respect to the basis given by the columns of a matrix B, we left-multiply the standard coordinates by B^{-1}
- To convert from coordinates with respect to the columns of B to standard coordinates, we left-multiply the coordinates by B
- The coordinate mapping from V to $\mathbb{R}^n$ which sends each vector to its coordinates with respect to some given (fixed) basis is called the *coordinate mapping*
- The coordinate mapping is linear and bijective, i.e., a vector space *isomorphism*
- Any linearly independent list of vectors in V can be extended to obtain a basis in V
- Any spanning list of vectors in V can be pruned to obtain a basis of V
- The *rank* of a linear transformation $T : V \to W$ is the dimension of the range of T
- The rank (or column rank) of an $m \times n$ matrix is dimension of the matrix's column space, and the *row rank* is the dimension of the matrix's row space (the span of the matrix's rows)
- The row rank of any matrix equals its column rank
- The rank of an $m \times n$ matrix A is equal to the smallest value of r such that A can be written as a product of an $m \times r$ matrix and an $r \times n$ matrix

Week 11 (3 April through 7 April)

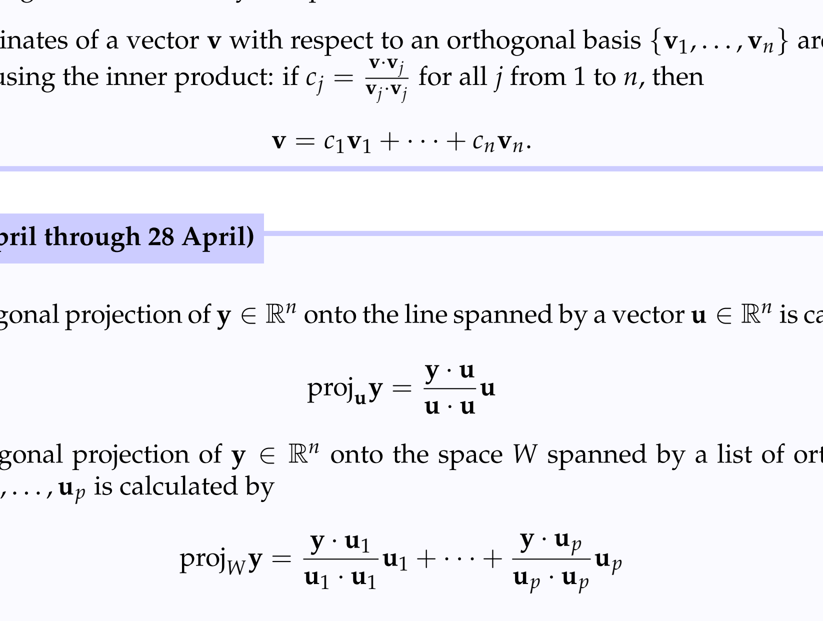
- If we have a vector represented via its coordinates with respect to some basis and we want its coordinates with respect to a different basis, then we left-multiply by the matix whose columns contain the coordinates of the old basis vectors with respect to the new ones
- The change of basis matrix from one basis $\mathcal{B}$ to another basis $\mathcal{C}$ is the inverse of the change of basis matrix from $\mathcal{C}$ back to $\mathcal{B}$
- An eigenvector $\mathbf{v}$ of an $n \times n$ matrix A is a vector with the property that $A\mathbf{v} = \lambda\mathbf{v}$ for some $\lambda \in \mathbb{R}$ (n other words, A maps $\mathbf{v}$ to a vector which is either zero or parallel to $\mathbf{v}$)
- The eigenspace associated with a given eigenvalue λ is the set of all solutions to the homogeneous sytsem $(A - \lambda I)\mathbf{v} = 0$, where I is the $n \times n$ identity matrix
- If a matrix has n linearly independent eigenvectors, then A's action on $\mathbb{R}^n$ can be understood simply. Each coordinate gets multiplied by the corresponding eigenvalue:
$$A(c_1\mathbf{v}_1 + \cdots c_n\mathbf{v}_n) = c_1\lambda_1\mathbf{v}_1 + c_n\lambda_n\mathbf{v}_n.$$
- This can be used to find a formula for the nth Fibonacci number in terms of the nth power of the golden ratio $\phi = \frac{1+\sqrt{5}}{2}$.

Week 12 (10 April through 14 April)

- Eigenvectors corresponding to distinct eigenvalues are linearly independent
- We can find eigenvalues by finding roots of the characteristic polynomial $\det(A - \lambda I)$
- An $n \times n$ matrix has at most n eigenvalues
- A matrix A is diagonalizable if $A = PDP^{-1}$ for some invertible P and diagonal D
- Diagonalizability is equivalent to having n linearly independent eigenvectors
- If some eigenvalues of A are repeated roots of the characteristic polynomial, there might not be a basis of $\mathbb{R}^n$ consisting of eigenvectors of A
- If two matrices A and B have the property that $A = PBP^{-1}$ for some invertible matrix P, then we say A and B are *similar*

Week 13 (17 April through 21 April)

- Complex numbers are expressions of the form $a + bi$ where a and b are real numbers (so real numbers count as complex too)
- The conjugate of $z = a + bi$ is $\overline{z} = a - bi$.
- The absolute value of $z = a + bi$ is $|z| = \sqrt{a^2 + b^2}$, and its argument is the angle between the positive x-axis and the pont (a, b)
- These expressions plotted in the plane (at (a, b)), and they support all the standard arithmetic (multiplication, addition, multiplicative inverses, and additive inverses), with the rule that we reduce i^2 to -1
- The characteristic polynomial of a matrix may have no real roots (such as the matrix of a 90-degree rotation in $\mathbb{R}^2$), but every degree-n polynomial does have a complex root
- So every $n \times n$ matrix A has a complex eigenvalue λ and a complex eigenvector (a vector $\mathbf{v}$ with n complex entries such that $A\mathbf{v} = \lambda\mathbf{v}$)
- If λ is a complex eigenvalue with complex eigenvector $\mathbf{v}$, then $\overline{\lambda}$ is a complex eigenvalue with eigenvector $\overline{\mathbf{v}}$.
- A matrix of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ represents a rotation with angle $\arg(a + bi)$ combined with a scaling by a factor of $\sqrt{a^2 + b^2}$
- Every 2×2 matrix with complex, non-real eigenvalues is similar to a scaling-rotation matrix (a matrix of the form described in the preceding item)
- The *inner product* of two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ is defined to be the scalar $\mathbf{u}^T\mathbf{v}$, also denoted $\mathbf{u} \cdot \mathbf{v}$ and called the *dot product*
- $\mathbf{u} \cdot \mathbf{u} = |\mathbf{u}|^2$
- $\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}||\mathbf{v}|\cos\theta$, where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$ (proof: law of cosines)
- $\mathbf{u} \cdot \mathbf{v} = 0$ if and only if $\mathbf{u}$ and $\mathbf{v}$ are orthogonal
- The dot product is linear: $\mathbf{u} \cdot (c\mathbf{v} + \mathbf{w}) = c\mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w})$
- A vector is called a unit vector if it has length 1
- If V is a subspace of $\mathbb{R}^n$, then the orthogonal complement $V^\perp$ is the set of vectors orthogonal to every vector in V
- $V^\perp$ is a subspace
- If A is an $m \times n$ matrix, then the orthogonal complement of the row space of A is the null space of A, and the orthogonal complement of the column space of A is the null space of A^T
- We can put this together into the following schematic diagram:



- A list of vectors in $\mathbb{R}^n$ is *orthogonal* if any pair of distinct vectors from the list is orthogonal
- Every orthogonal list is linearly independent
- The coordinates of a vector $\mathbf{v}$ with respect to an orthogonal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ are easy to compute using the inner product: if $c_j = \frac{\mathbf{v} \cdot \mathbf{v}_j}{\mathbf{v}_j \cdot \mathbf{v}_j}$ for all j from 1 to n, then
$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n.$$

Week 14 (24 April through 28 April)

- The orthogonal projection of $\mathbf{y} \in \mathbb{R}^n$ onto the line spanned by a vector $\mathbf{u} \in \mathbb{R}^n$ is calculated by
$$\text{proj}_{\mathbf{u}}\mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}\mathbf{u}$$
- The orthogonal projection of $\mathbf{y} \in \mathbb{R}^n$ onto the space W spanned by a list of orthogonal vectors $\mathbf{u}_1, \ldots, \mathbf{u}_p$ is calculated by
$$\text{proj}_W\mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p}\mathbf{u}_p$$
- An orthonormal list is an orthogonal list all of whose vectors have length 1
- A matrix whose columns are orthonormal is called an *orthogonal* matrix (yes, this is confusing)
- The orthogonal projection of $\mathbf{y}$ onto $\mathbf{W}$ is the point in W which is nearest to $\mathbf{y}$
- If U is an orthogonal $n \times p$ matrix, then UU^T is the matrix of the transformation which projects each vector in $\mathbb{R}^n$ onto the p-dimensional subspace of $\mathbb{R}^n$ spanned by the columns of U
- To come up with a list of vectors $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p$ which as the same span as some specified linearly independent list $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$, we successively subtract off orthogonal projections:
$$\begin{aligned} \mathbf{b}_1 &= \mathbf{v}_1 \\ \mathbf{b}_2 &= \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{b}_1}{\mathbf{b}_1 \cdot \mathbf{b}_1}\mathbf{b}_1 \\ \mathbf{b}_3 &= \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{b}_1}{\mathbf{b}_1 \cdot \mathbf{b}_1}\mathbf{b}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{b}_2}{\mathbf{b}_2 \cdot \mathbf{b}_2}\mathbf{b}_2 \end{aligned}$$
etc.
- The least squares problem in statistics asks for the slope and y-intercept of the line which "best fits" a collection of data points which are ordered pairs $\{(x_i, y_i)\}_{i=1}^n$, in the sense that the sum of squared errors
$$(y_1 - \alpha - \beta x_1)^2 + (y_2 - \alpha - \beta x_2)^2 + \cdots + (y_n - \alpha - \beta x_n)^2$$
is minimized
- We let $A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$, so we're looking to minimize $|A\mathbf{x} - \mathbf{b}|^2$
- To minimize $|A\mathbf{x} - \mathbf{b}|^2$, we find the orthogonal projection $\widehat{\mathbf{b}}$ of $\mathbf{b}$ onto the column space of A and choose a value of $\mathbf{x}$ satisfying $A\mathbf{x} = \widehat{\mathbf{b}}$
- $\widehat{\mathbf{b}}$ is characterized by $A^T(\mathbf{b} - \widehat{\mathbf{b}}) = \mathbf{0}$
- The solution of this equation is $\mathbf{x} = (A^TA)^{-1}A^T\mathbf{b}$ if A^TA is invertible