MATH 118, Spring 2020, Linear Algebra Key Ideas

Taken in part from
Introduction to Linear Algebra, 4e,
Gilbert Strang

Notes compiled by Bobby McDonald
Yale University, 2020

Contents

Chapter 1

Introduction to Vectors

1.1 Vectors and linear combinations

1.1. Key Ideas

- A vector $\mathbf{v}$ in two-dimensional space has two components v_1 and v_2.
- $\mathbf{v} + \mathbf{w} = \langle v_1 + w_1, v_2 + w_2 \rangle$ and $c\mathbf{v} = \langle cv_1, cv_2 \rangle$ are found a component at a time.
- A linear combination of three vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ is $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$.
- In three dimensions, *all* linear combinations of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ typically fill a line, then a plane, then the whole space $\mathbb{R}^3$.

1.2 Lengths and dot products

1.2. Key Ideas

- The dot product $\mathbf{v} \cdot \mathbf{w}$ multiplies each component v_i by w_i and adds all $v_i w_i$.
- The length $\|\mathbf{v}\|$ of a vector $\mathbf{v}$ is the square root of $\mathbf{v} \cdot \mathbf{v}$.
- $\mathbf{u} = \mathbf{v}/\|\mathbf{v}\|$ is a **unit vector.** Its length is 1
- The dot product $\mathbf{v} \cdot \mathbf{w} = 0$ when the vectors $\mathbf{v}$ and $\mathbf{w}$ are perpendicular
- If θ is the angle between $\mathbf{v}$ and $\mathbf{w}$, then $\cos\theta = \dfrac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|}$

1.3 Matrices

1.3. Key Ideas

- **Matrix times vector:** $A\mathbf{x} =$ linear combination of the columns of A with x_i as weights.
- The solution to $A\mathbf{x} = \mathbf{b}$ is $\mathbf{x} = A^{-1}\mathbf{b}$, when A is invertible.
- The difference matrix A is inverted by the sum matrix $S = A^{-1}$.
- The cyclic matrix C has no inverse. Its three columns lie in the same plane. Those dependent columns add to the zero vector. $C\mathbf{x} = \mathbf{0}$ has many solutions.

Chapter 2

Solving Linear Equations

2.1 Vectors and linear equations

2.1. Key Ideas

- The basic operations on vectors are multiplication $c\mathbf{v}$ and vector addition $\mathbf{v} + \mathbf{w}$.
- Together those operations give linear combinations $c\mathbf{v} + d\mathbf{w}$.
- Matrix-vector multiplication $A\mathbf{x}$ can be computed by dot products, a row at a time, but $A\mathbf{x}$ should be understood as a linear combination of the columns of A.
- $A\mathbf{x} = \mathbf{b}$ asks for a linear combination of the columns of A that produces $\mathbf{b}$
- Each equation $A\mathbf{x} = \mathbf{b}$ gives a line ($n = 2$), plane ($n = 3$), or "hyperplane" ($n > 3$). They intersect at the solution or solutions, if any.

2.2 The idea of elimination

2.2. Key Ideas

- A linear system becomes upper triangular after elimination.
- We subtract ℓ_{ij} times equation j from equation i to make the (i, j) entry zero, where

$$\ell_{ij} = \frac{(i, j) \text{ entry}}{\text{pivot in row } j}.$$

- A zero in the pivot position can be repaired if there is a nonzero below it.
- The upper triangular system is solved by back substitution.
- When breakdown is permanent, the system has no solution or infinitely many.

2.3 Elimination using matrices

2.3. Key Ideas

- $A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots x_n\mathbf{a}_n$
- row reduction and echelon form
- row reduction can be interpretted as products of elementary matrices with A
- When A multiplies any matrix B, it multiplies each column of B

2.4 Rules for matrix operations

2.4. Key Ideas

- The (i, j) entry of AB is the dot product of row i of A with column j of B.
- An $m \times n$ matrix times an $n \times p$ matrix gives an $m \times p$ matrix, and uses mnp separate multiplications.
- $A(BC) = (AB)C$, but $AB \neq BA$ in general
- Block multiplication is allowed when the block shapes match correctly
- Block elimination produces the *Schur Complement* $D - CA^{-1}B$.

2.5 Invers matrices

2.5. Key Ideas

- The inverse matrix gives $AA^{-1} = I$ and $A^{-1}A = I$.
- A is invertible if and only if it has n pivots
- If $A\mathbf{x} = 0$ for a nonzero vector $\mathbf{x}$, then A has no inverse
- The inverse of AB is $B^{-1}A^{-1}$, and $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$.
- Reducing $[\ A \quad I\ ]$ to reduced row echelon form gives $[\ I \quad A^{-1}\ ]$.

2.6 LU factorization

2.6. Key Ideas

- Elimination (with now row exchanges) factors A into LU.
- L contains the numbers ℓ_{ij} that muiltiply pivot rows to get from A to U.
- On the right side we solve $L\mathbf{c} = \mathbf{b}$ (forward) and $U\mathbf{x} = \mathbf{b}$ (backward)

2.7 Transposes and permutations

2.7. Key Ideas

- The rows of A are the columns of A^T.
- The transpose of AB is $B^T A^T$, and $(A^T)^{-1}$ is $(A^{-1})^T$.
- The dot product is $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y}$.
- When A is symmetric ($A^T = A$), it's LDU factorization is symmetric $A = LDL^T$.
- A permutation matrix P has a 1 in each row and column, and $P^T = P^{-1}$.
- If A is invertible then a permutation P will reorder its rows for $PA = LU$.

Chapter 3

Vector Spaces and Subspaces

3.1 Spaces of vectors

3.1. Key Ideas

- $\mathbb{R}^n$ contains all column vectors with n real components.
- $M_{2\times 2}$, $\mathbf{F}$, and $\{\mathbf{0}\}$ are vector spaces.
- A subspace containing $\mathbf{v}$ and $\mathbf{w}$ must contain all linear combinations $c\mathbf{v} + d\mathbf{w}$.
- The combinations of the columns of A form the column space $\mathrm{Col}(A)$. The column space is "spanned" by the columns.
- $A\mathbf{x} = \mathbf{b}$ has a solution exactly when $\mathbf{b}$ is in the column space of A.

3.2 The null space, solutions to $A\mathbf{x} = \mathbf{0}$

3.2. Key Ideas

- The null space $\mathrm{Nul}(A)$ is a subspace of $\mathbb{R}^n$. It contains all solutions to $A\mathbf{x} = \mathbf{0}$.
- Elimination produces an echelon matrix U, and then a row reduced R, with pivot columns and free columns.
- Every free column of U or R leads to a special solution. The free variable equals 1 and the other free variables equal 0. Back substitution solves $A\mathbf{x} = \mathbf{0}$.
- The complete solution to $A\mathbf{x} = \mathbf{0}$ is the linear combination of all the special solutions.
- If $n > m$, then A has at least one column without pivots, giving a special solution. So there are nonzero vectors in $\mathrm{Nul}(A)$.

3.3 The complete solution to $A\mathbf{x} = \mathbf{b}$

3.3. Key Ideas

- The rank r is the number of pivots. The matrix R has $m - r$ zero rows.
- $A\mathbf{x} = \mathbf{b}$ is solvable if and only if the last $m - r$ equations reduce to $0 = 0$.
- One particular solution $\mathbf{x}_p$ has all free variables equal to zero.
- The pivot variables are determined after the free variables are chosen.
- Full column rank $r = n$ means no free variables: one solution or none.
- Full row rank $r = m$ means one solution $m = n$ or infinintely many if $m < n$.

3.4 Independence, basis, and dimension

3.4. Key Ideas

- The columns of A are independint if $\mathbf{x} = 0$ is the only solution to $A\mathbf{x} = \mathbf{0}$.
- The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ span a space if their combinations fill that space.
- A basis consists of linearly independent vectors that span the space, and every vector is a unique combination of vectors in that basis.
- All bases for a space have the same number of vectors, this number is called the dimension.
- The pivot columns are one basis for the column space. The dimension is $\mathbf{r}$.

3.5 Dimensions of the four subspaces

3.5. Key Ideas

- The four subspaces are the row space, column space, nullspace, and left nullspace.
- The r pivot rows of r are a basis for the row spaces of R and A
- The r pivot columns of A are a basis for the column space.
- The $n - r$ special solutions are a basis for the null spaces of A and R.
- The last $m - r$ rows of I are a basis for the left null space of R.
- The last $m - r$ rows of E are a basis for the left null space of A.

Chapter 4

Orthogonality

4.1 Orthogonality of the four subspaces

4.1. Key Ideas

- Subspaces V and W are orthogonal if every $\mathbf{v}$ in V is orthogonal to every $\mathbf{w}$ in W.
- V and W are "orthogonal complements" if W contains all vectors perpendicular to V (and vice versa). Inside $\mathbf{R}^n$, the dimensions of the complements V and W add to n.
- The nullspace and rowspace are orthogonal complements, as are the column and left null spaces.
- Any n linearly independent vectors in $\mathbb{R}^n$ will span $\mathbb{R}^n$.
- Every $\mathbf{x}$ in $\mathbb{R}^n$ has a nullspace component $\mathbf{x}_n$ and a row space component $\mathbf{x}_r$.

4.2 Projections

4.2. Key Ideas

- The projection of $\mathbf{b}$ onto the line through $\mathbf{a}$ is $\mathbf{p} = \mathbf{a}\widehat{\mathbf{x}} = \dfrac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{a} \cdot \mathbf{a}}\mathbf{a}$
- Projecting $\mathbf{b}$ onto a subsspace leaves $\mathbf{z} = \mathbf{b} - \widehat{\mathbf{p}}$ perpendicular to the subspace
- When A has full rank n, the equation $A^T A\widehat{\mathbf{x}} = A^T\mathbf{b}$ leads to $\widehat{\mathbf{x}}$ and $\mathbf{p} = A\mathbf{x}$.
- The projection matrix $P = A(A^T A)^{-1}A^T$ has $P^T = P$ and $P^2 = P$.

4.3 Least squares

4.3. Key Ideas

- The least squares solution $\widehat{\mathbf{x}}$ minimizes $E = \|A\mathbf{x} - \mathbf{b}\|^2$. This is the sum of squares of the errors in the m equations ($m > n$).
- The best $\widehat{\mathbf{x}}$ comes from the normal equations $A^T A\mathbf{x} = A^T\mathbf{b}$.
- To fit m points by a line $b = C + Dt$, the normal equations give C and D.
- The heights of the best line are $\mathbf{p} = (p_1, \ldots, p_m)$. The vertical distances to the data points are the errors $\mathbf{e} = (e_1, \ldots, e_m)$.
- If we try to fit m points by a combination of $n < m$ functions, the m equations $A\mathbf{x} = \mathbf{b}$ are generally unsolvable. The n equations $A^T A\widehat{\mathbf{x}} = A^T\mathbf{b}$ give the least squares solution – the ombination with the smallest mean square error.

4.4 Orthogonal bases and Gram-Schmidt

4.4. Key Ideas

- If the orthonormal vectors $\mathbf{q}_1, \ldots, \mathbf{q}_n$ are the columns of Q, then $\mathbf{q}_i \cdot \mathbf{q}_j = 0$ and $\mathbf{q}_i\mathbf{q}_i = 1$. In other words, $Q^T Q = I$.
- If Q is square then Q is called an orthogonal matrix, and $Q^T = Q^{-1}$
- The length of $Q\mathbf{x}$ equals the length of $\mathbf{x}$: $\|Q\mathbf{x}\| = \|\mathbf{x}\|$
- The projection onto the column space spanned by the $\mathbf{q}$ is $P = QQ^T$.
- If Q is square, then $P = I$, and every $\mathbf{b} = (\mathbf{b} \cdot \mathbf{q}_1)\mathbf{q}_1 + (\mathbf{b} \cdot \mathbf{q}_2)\mathbf{q}_2 + \cdots + (\mathbf{b} \cdot \mathbf{q}_n)\mathbf{q}_n$
- Gram-Schmidt produces orthonormal vectors $\mathbf{q}_1$, $\mathbf{q}_2$, and $\mathbf{q}_3$ from linearly independent $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$. In matrix form, this is the factorization $A = QR$.