CSE 6643 Homework 1

Karl Hiner

1   Basics [25 pts]

(a) [5 pts]

Suppose that $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4$ is a basis of the vector space $V \subset \mathbb{R}^n$. Prove that the list

$$\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{v}_2 + \boldsymbol{v}_3, \boldsymbol{v}_3 + \boldsymbol{v}_4, \boldsymbol{v}_4 \tag{1}$$

is also a basis of V.

We can prove these vectors form a basis of V if we show that they are linearly independent and span the vector space.

(a1) Linear independence:

To show that the given vectors are linearly independent, we first assume the vectors are linearly *dependent*, and find the coefficients for the linear combination that equals the zero vector. If the only solution is that the coefficients are all zero, then the vectors are linearly independent.

Suppose there exists $\{a_1, a_2, a_3, a_4\}, a_i \in \mathbb{R}, a_i \neq 0$ such that

$$a_1(\boldsymbol{v}_1 + \boldsymbol{v}_2) + a_2(\boldsymbol{v}_2 + \boldsymbol{v}_3) + a_3(\boldsymbol{v}_3 + \boldsymbol{v}_4) + a_4\boldsymbol{v}_4 = \boldsymbol{0}. \tag{2}$$

Rearranging terms,

$$a_1\boldsymbol{v}_1 + (a_1 + a_2)\boldsymbol{v}_2 + (a_2 + a_3)\boldsymbol{v}_3 + (a_3 + a_4)\boldsymbol{v}_4 = \boldsymbol{0}. \tag{3}$$

Since $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4$ is a basis of V, we can rewrite this as

$$a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + a_3\boldsymbol{v}_3 + a_4\boldsymbol{v}_4 = \boldsymbol{0}. \tag{4}$$

Since $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4$ are linearly independent, we can conclude that $a_1 = a_2 = a_3 = a_4 = 0$.

Thus, the vectors $\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{v}_2 + \boldsymbol{v}_3, \boldsymbol{v}_3 + \boldsymbol{v}_4, \boldsymbol{v}_4$ are linearly independent.

(a2) Span:

Since $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4)$ is a basis of V, any vector $\boldsymbol{u} \in V$ can be expressed as a linear combination with coefficients $\{a_1, a_2, a_3, a_4\}, a_i \in \mathbb{R}$.

If we can show that each vector in $L = (\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{v}_2 + \boldsymbol{v}_3, \boldsymbol{v}_3 + \boldsymbol{v}_4, \boldsymbol{v}_4)$ can be expressed in terms of $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4)$, then we will show that any vector $\boldsymbol{u} \in V$ can be also be expressed as a linear combination of the vectors in L:

$$\boldsymbol{v}_1 + \boldsymbol{v}_2 = (1)\boldsymbol{v}_1 + (1)\boldsymbol{v}_2 \tag{5}$$
$$\boldsymbol{v}_2 + \boldsymbol{v}_3 = (1)\boldsymbol{v}_2 + (1)\boldsymbol{v}_3 \tag{6}$$
$$\boldsymbol{v}_3 + \boldsymbol{v}_4 = (1)\boldsymbol{v}_3 + (1)\boldsymbol{v}_4 \tag{7}$$
$$\boldsymbol{v}_4 = (1)\boldsymbol{v}_4 \tag{8}$$

Thus, L spans V.

Therefore, the list $\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{v}_2 + \boldsymbol{v}_3, \boldsymbol{v}_3 + \boldsymbol{v}_4, \boldsymbol{v}_4$ is also a basis of V.

(b) [10 pts]

For U a subspace of the vector space $V \subset \mathbb{R}^n$ with $\dim(U) = \dim(V)$. Prove that $U = V$.

Since we are given that U is a subspace of V, every vector in U is also in V, so U is a subset of V.

Since $\dim(U) = \dim(V)$, U and V have the same number of linearly independent vectors. This, combined with the previously shown result that every vector in U is also in V, implies that U and V have the same set of linearly independent vectors, and thus contain the same set of vectors.

(c) [10 pts]

Show that the subspaces of $\mathbb{R}^3$ are precisely $\{0\}$, $\mathbb{R}^3$, all lines in $\mathbb{R}^3$ through the origin, and all planes in $\mathbb{R}^3$ through the origin.

First, we must show that these sets are indeed subspaces of $\mathbb{R}^3$. For each set $\mathbb{U}$, we must show that the following conditions hold in order to show that it is a subspace of $\mathbb{R}^3$:

1. $\{0\} \in \mathbb{U}$

2. Closed under addition: $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{U} \implies \boldsymbol{u} + \boldsymbol{v} \in \mathbb{U}$

3. Closed under scalar multiplication: $a \in \mathbb{R}, \boldsymbol{u} \in \mathbb{U} \implies a\boldsymbol{u} \in \mathbb{U}$

Enumerating each of the four sets:

1. $\mathbb{U} := \{0\}$: Since $\{0\} \in \{0\}$, and $dim(\{0\}) = 0$ (it is the only element in the set), it is trivially a subspace of R^3.

2. $\mathbb{U} :=$ All lines in $\mathbb{R}^3$ through the origin:

   - Every line in $\mathbb{R}^3$ is defined by the set of points $\boldsymbol{u} = \boldsymbol{b} + c\boldsymbol{v}$, with $c \in \mathbb{R}, \boldsymbol{v} \neq \{0\}$. Since the set only includes lines passing through the origin, $\boldsymbol{u} = c\boldsymbol{v}$.

   - $\{0\} \in \mathbb{U}$ by construction.

   - For any $\boldsymbol{u} = a\boldsymbol{u}'$ and $\boldsymbol{v} = b\boldsymbol{v}'$, with $\boldsymbol{u}', \boldsymbol{v}' \in R^3, \boldsymbol{u}', \boldsymbol{v}' \neq \{0\}$ and $a, b \in \mathbb{R}$,

   $$c\boldsymbol{u} + d\boldsymbol{v} = ca\boldsymbol{u}' + db\boldsymbol{v}' \tag{9}$$
   $$= ca(\boldsymbol{u}' + \frac{ca}{db}\boldsymbol{v}') \tag{10}$$
   $$= c'\boldsymbol{w} \qquad (c' \equiv (ca) \in \mathbb{R}, \boldsymbol{w} \equiv (\boldsymbol{u}' + \frac{ca}{db}\boldsymbol{v}') \in \mathbb{R}^3) \tag{11}$$

   Thus, this set is closed under linear combinations (satisfying conditions 2 and 3).

3. $\mathbb{U} :=$ All planes in $\mathbb{R}^3$ through the origin: We could prove this algebraically using a similar approach to the case of lines through the origin. However, I will avoid repetition by noting that we can extend our proof over lines in $\mathbb{R}^3$ to a proof over planes in $\mathbb{R}^3$ by replacing our scalars (c, d, a, b) with 3-vectors (plane normals), without violating linearity. We could also make a geometric argument that multiplying a plane through the origin by a constant creates a new plane through the origin, as does adding together two planes through the origin.

4. $\mathbb{U} := \mathbb{R}^3$: Since $\mathbb{R}^3$ is a space, and every space is its own subspace, $\mathbb{R}^3$ is a subspace of $\mathbb{R}^3$.

Finally, we must show that any subspace in $\mathbb{R}^3$ must itself be a subspace of one of these four subspaces.

First, note that there can exist no subspace $V \subset \mathbb{R}^n$ with $\dim(V) > n$, since any V with $\dim(V) > n$ must have more than n basis vectors, and thus will necessarily have vectors that are not contained in $\mathbb{R}^n$. Combining this with the result of b above, which holds that for any subspace U of $V \subset \mathbb{R}^n$ with $\dim(U) = \dim(V), U = V$, we can conclude that there must be exactly $n + 1$ unique subspaces of $V \subset \mathbb{R}^n$ - one for each dimension $0, 1, ..., n$.

In this case, $V = \mathbb{R}^3$, so $n = 3$, and thus there are exactly 4 subspaces of $\mathbb{R}^3$. Since we have enumerated proofs for each of these four cases above, we are done.

2 Norm Equivalencies [25 pts]

In a finite-dimensional space, all norms are equivalent. In this problem, you will be asked to verify this theorem for some special norms. Prove the following inequalities.

Let $\boldsymbol{x} \in \mathbb{R}^n$ be an n-dimensional vector. Let $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ be an $m \times n$ matrix. Then:

(a) [10 pts]

$$\|\boldsymbol{x}\|_\infty \leq \|\boldsymbol{x}\|_2 \leq \sqrt{n}\|\boldsymbol{x}\|_\infty.$$

Proof:

$$\|\boldsymbol{x}\|_\infty \leq \qquad \|\boldsymbol{x}\|_2 \leq \qquad \sqrt{n}\|\boldsymbol{x}\|_\infty \tag{12}$$

$$\max_i |x_i| \leq \qquad \sqrt{\sum_{i=1}^n x_i{}^2} \leq \qquad \sqrt{n}\max_i |x_i| \quad \text{(expanding definitions)} \tag{13}$$

$$(\max_i |x_i|)^2 \leq \qquad \sum_{i=1}^n x_i{}^2 \leq \qquad n(\max_i |x_i|)^2 \quad \text{(squaring all (positive) terms)} \tag{14}$$

$$x_m{}^2 \leq \qquad \sum_{i=1}^n x_i{}^2 \leq \qquad nx_m{}^2 \quad \text{(let } x_m \equiv \max_i |x_i|) \tag{15}$$

Consider the extreme case where $x_i \equiv x_m, \forall x_i \in \boldsymbol{x}$. That is, all elements in $\boldsymbol{x}$ are the same element, and so its sum-of-squares is maximally large relative to its maximum absolute element, x_m. Then,

$$x_m{}^2 \leq \qquad \sum_{i=1}^n x_m{}^2 \leq \qquad nx_m{}^2 \quad \text{(subsituting } x_i = x_m) \tag{16}$$

$$x_m{}^2 \leq \qquad nx_m{}^2 \leq \qquad nx_m{}^2 \quad \text{(QED for this case)} \tag{17}$$

Now, consider the other extreme, where $x_1 \neq 0, x_i \equiv 0, \forall i \neq 1$. That is, all elements in $\boldsymbol{x}$ are 0, except a single non-zero element, arbitrarily chosen to be at index $i = 1$. So, the absolute maximum element $x_m = |x_i|$, and the sum-of-squares of $\boldsymbol{x}$ is maximally small relative to x_m. Then,

$$x_m{}^2 \leq \qquad \sum_{i=1}^{n} x_m{}^2 \leq \qquad nx_m{}^2 \tag{18}$$

$$x_m{}^2 \leq \qquad x_m{}^2 \leq \qquad nx_m{}^2 \quad \text{(QED for this case)} \tag{19}$$

All other cases will result in a $\sum_{i=1}^{n} x_i{}^2$ between these extremes.

QED.

(b) [7.5 pts]

$$\|\boldsymbol{A}\|_\infty \leq \sqrt{n}\|\boldsymbol{A}\|_2.$$

NOTE: I used ChatGPT to help me with this and the next problem (c), and (to a lesser extent) the "Perturbing" section. I spent a great deal of time working with my answers, but don't fully understand this material enough to have generated such a proof completely on my own. It's been years since I've taken a full linear algebra course, and I am working very hard to build up my understanding in tandem with this course. I'm confident that I will be able to catch up, but this early in the course I have to admit I feel a bit behind in this theoretical material!

$$\|\boldsymbol{A}\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^{n} |\boldsymbol{A}_{i,j}| \qquad \text{(def. of matrix infinity-norm)} \tag{20}$$

$$\leq \sqrt{n} \max_{1 \leq i \leq m} \sqrt{\sum_{j=1}^{n} |\boldsymbol{A}_{i,j}|^2} \qquad \text{(Cauchy-Schwarz inequality)} \tag{21}$$

$$\leq \sqrt{n} \max_{1 \leq i \leq m} \|\boldsymbol{A}_i\|_2 \qquad \text{(def. of vector L2 norm)} \tag{22}$$

$$= \sqrt{n}\|\boldsymbol{A}\|_2 \qquad \text{(def. of matrix 2-norm)} \tag{23}$$

(c) [7.5 pts]

$$\|\boldsymbol{A}\|_2 \leq \sqrt{m}\|\boldsymbol{A}\|_\infty.$$

$$\|\boldsymbol{A}\|_2 = \sqrt{\lambda_{max}(\boldsymbol{A}^T\boldsymbol{A})} \qquad \text{(def. of 2-norm)} \tag{24}$$

$$\leq \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2} \qquad \text{(def. of 2-norm)} \tag{25}$$

$$\leq \sqrt{m} \max_{1 \leq j \leq n} \left(\sum_{i=1}^{m} |a_{ij}|\right) \qquad \text{(Cauchy-Schwarz inequality)} \tag{26}$$

$$= \sqrt{m}|\boldsymbol{A}|_\infty \qquad \text{(def. of infinity norm)} \tag{27}$$

4

3   Perturbing [25 pts]

For $u, v \in \mathbb{K}^m$, the matrix $A := I + uv^*$ is called a *rank-one* perturbation of the identity.

(a) [15 pts]

Show that if A is nonsingular, then its inverse has the form $A^{-1} = I + \alpha uv^*$ for some scalar α, and give an expression for α.

- Multiply both sides by A to get $AA^{-1} = A(I + \alpha uv^*)$.

- Since A is nonsingular, it has an inverse, so we can substitute $AA^{-1} = I$.

- This gives us $I = A(I + \alpha uv^*)$.

- Substituting $A = I + uv^*$, $I = I + \alpha uv^* + uv^*$.

- Solving for α, we get $\alpha = -1$.

Therefore, the inverse of $A = I + uv^*$ is $A^{-1} = I + \alpha uv^* = I - uv^*$.

(b) [5 pts]

For what u and v is A nonsingular?

For A to be nonsingular, u and v must be linearly independent, so that $u \cdot v \neq 0$. This guarantees that A has a non-zero determinant, and is thus invertible.

(c) [5 pts]

If A is singular, what is null(A)?

If A is singular, then u and v are *not* linearly independent, and so $Ax = 0$ has non-trivial solutions.

Specifically, let's assume we can express u as a scalar multiple of v, so that $u = av$, for some nonzero scalar a. Then,

$A = I + uv^* = I + avv^*$

In this case, the null space will be the one-dimensional subspace spanned by v.

Julia (e) Optimization [5 pts]

Make sure that your code does not allocate any memory, as evidenced by the `@btime` calls in `HW1_driver.jl` returning (`0 allocations: 0 bytes`). This is important for performance reasons since allocating memory may be orders of magnitudes slower than floating-point arithmetic. Now try reordering the for-loops in your implementations from parts (a) and (b) and observe the resulting timings provided by `@btime`. Which order leads to the best and worst performance? What are the corresponding wall-clock times as measured by `@btime`?

Julia (e) Optimization: `matvec`

For `matvec`, iterating over rows in the outer loop and columns in the inner loop resulted in:

```
Timing the matvec function...
1.685 ms (0 allocations: 0 bytes)
```

Switching the loop order to columns-first resulted in a major speedup:

```
Timing the matvec function...
274.375 microseconds (0 allocations: 0 bytes)
```

Julia(e) Optimization: `matmul`

For `matmul`, here are my timing experiments, ordered by fastest to slowest, where e.g. B_i, B_j, C_j means iterating first over B's rows, then B's columns, then C's columns as the innermost loop.

- C_j, B_j, B_i: 892.252 ms
- B_j, C_j, B_i: 1.117 s
- C_j, B_i, B_j: 1.757 s
- B_i, C_j, B_j: 1.772 s
- B_j, B_i, C_j: 2.628 s
- B_i, B_j, C_j: 2.736 s