MA 8463 - Homework 1

Jacob Kutch

1. Prove $\|\cdot\|_\infty = \lim_{p\to\infty} \|\cdot\|_p$

a. Verify $\lim_{p\to\infty} (1+x^p)^{1/p} = 1, \ \forall |x| \le 1$

$|x| \le 1 \Rightarrow -1 \le x \le 1$; Using Squeeze Theorem:

$$-1 \le x \le 1 \Rightarrow (-1)^p \le x^p \le 1 \Rightarrow (-1)^p + 1 \le x^p + 1 \le 2$$

$$\Rightarrow ((-1)^p + 1)^{1/p} \le (x^p+1)^{1/p} \le 2^{1/p}$$

$$\lim_{p\to\infty} ((-1)^p+1)^{1/p} \le \lim_{p\to\infty} (x^p+1)^{1/p} \le \lim_{p\to\infty} (2^{1/p})$$

*Originally, I tried

$$\lim_{p\to\infty} \left(e^{\ln\left(((-1)^p+1)^{1/p}\right)}\right) \le \lim_{p\to\infty} (x^p+1)^{1/p} \le 1$$

$$\lim_{p\to\infty} \left(e^{\frac{1}{p}\ln((-1)^p+1)}\right) \le \lim_{p\to\infty} (x^p+1)^{1/p} \le 1$$

$$e^0 \le \lim_{p\to\infty} (x^p+1)^{1/p} \le 1 \Rightarrow 1 \le \lim_{p\to\infty} (x^p+1)^{1/p} \le 1$$

I now see the issue in that the sequence $\{((-1)^p+1)^{1/p}\}$ is oscillatory between 0 and 1 and squeeze theorem only holds for even p.

1. b. Let $x = (a,b)^T$ with $|a| \geq |b|$.

Prove $\lim_{p \to \infty} \|x\|_p = |a|$

Proof by more squeeze theorem:

$$\lim_{p \to \infty} \left(\sum_i |x_i|^p \right)^{\frac{1}{p}} = \lim_{p \to \infty} \left(|a|^p + |b|^p \right)^{\frac{1}{p}}$$

$$0 \leq |b| \leq |a|$$

$$0 \leq |b|^p \leq |a|^p$$

$$|a|^p \leq |a|^p + |b|^p \leq 2|a|^p$$

$$\left(|a|^p\right)^{\frac{1}{p}} \leq \left(|a|^p + |b|^p\right)^{\frac{1}{p}} \leq \left(2|a|^p\right)^{\frac{1}{p}}$$

$$|a| \leq \lim_{p \to \infty} \left(|a|^p + |b|^p\right)^{\frac{1}{p}} \leq \lim_{p \to \infty} 2^{\frac{1}{p}} |a|$$

$$|a| \leq \lim_{p \to \infty} \left(|a|^p + |b|^p\right)^{\frac{1}{p}} \leq |a|.$$

So $\lim_{p \to \infty} \left(|a|^p + |b|^p\right)^{\frac{1}{p}} = |a|$

$$\Rightarrow \lim_{p \to \infty} \|x\|_p = |a|$$

1.c. Prove $\lim_{p\to\infty} \|x\|_p = \|x\|_\infty$ for $x \in \mathbb{R}^n$

Proof by even more squeeze theorem:

$$\lim_{p\to\infty} \|x\|_p = \lim_{p\to\infty} \left(\sum_i |x_i|^p\right)^{\frac{1}{p}}$$

Consider the smallest case of $\sum_i |x_i|^p = (\max_i |x_i|)^p$ where $\max_i |x_i|$ is the only nonzero element of x, as well as the largest case of $\sum_i |x_i|^p = n(\max_i |x_i|)^p$ where all elements of x are equal to $\max_i |x_i|$ for all n elements of x. Thus,

$$(\max_i |x_i|)^p \leq \sum_i |x_i|^p \leq n(\max_i |x_i|)^p$$

$$\lim_{p\to\infty}\left((\max_i |x_i|)^p\right)^{\frac{1}{p}} \leq \lim_{p\to\infty}\left(\sum_i |x_i|^p\right)^{\frac{1}{p}} \leq \lim_{p\to\infty}\left(n(\max_i |x_i|)^p\right)^{\frac{1}{p}}$$

$$\max_i |x_i| \leq \lim_{p\to\infty}\left(\sum_i |x_i|^p\right)^{\frac{1}{p}} \leq \underbrace{\lim_{p\to\infty} n^{\frac{1}{p}}}_{1}(\max_i |x_i|)$$

$$\max_i |x_i| \leq \lim_{p\to\infty}\left(\sum_i |x_i|^p\right)^{\frac{1}{p}} \leq \max_i |x_i|$$

Thus $\lim_{p\to\infty}\left(\sum_i |x_i|^p\right)^{\frac{1}{p}} = \max_i |x_i| = \|x\|_\infty$

$$\Rightarrow \lim_{p\to\infty} \|x\|_p = \|x\|_\infty$$

2. Let $A \in \mathbb{R}^{n \times n}$ be positive definite with Cholesky factor R so that $A = R^T R$.

a. Verify $\|x\|_A = \|Rx\|_2 \ \forall x \in \mathbb{R}^n$

$$\Rightarrow (x^TAx)^{\frac{1}{2}} = \|Rx\|_2$$

$$(x^TAx)^{\frac{1}{2}} = (x^TR^TRx)^{\frac{1}{2}} = ((Rx)^TRx)^{\frac{1}{2}}$$

For clarity, let $u = Rx$. Since $Rx \in \mathbb{R}^n$,

$(Rx)^TRx = u^Tu$ is the dot product of u and itself.

Thus $(u^Tu)^{\frac{1}{2}} = \|u\|_2$.

So $(x^TAx)^{\frac{1}{2}} = \|Rx\|_2 \Rightarrow \|x\|_A = \|Rx\|_2 \ \forall x \in \mathbb{R}^n$.

b. Prove that $\|\cdot\|_2$ being a norm on $\mathbb{R}^n$ implies that $\|\cdot\|_A$ is a norm on $\mathbb{R}^n$.

For $\|x\|_A$ for $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times n}$ is necessarily positive definite and may be decomposed into its Cholesky factorization $A = R^TR$

Reversing steps in (a), any $\ell 2$-norm $\|u\|_2$ of $u \in \mathbb{R}^n$ can be written instead as $(u \cdot u)^{\frac{1}{2}} = (u^Tu)^{\frac{1}{2}}$. Considering $u = Rx \ \forall x \in \mathbb{R}^n$,

$$((Rx)^T(Rx))^{\frac{1}{2}} = (x^TR^TRx)^{\frac{1}{2}} = (x^TAx)^{\frac{1}{2}} = \|x\|_A.$$

Since $\|x\|_A$ is therefore an $\ell 2$-norm of Rx, it also a norm on $\mathbb{R}^n$

3. Prove that $\forall x \in \mathbb{R}^n$,

$$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2 \leq n\|x\|_\infty$$

$$\max_i |x_i| \leq \left(\sum_i^n |x_i|^2\right)^{\frac{1}{2}} \leq \sum_i^n |x_i| \leq \sqrt{n}\left(\sum_i^n |x_i|^2\right)^{\frac{1}{2}} \leq n\left(\max_i |x_i|\right)$$

Consider $\|x\|_2^2 = \sum_{i=1}^n |x_i|^2$. By Cauchy-Schwarz Inequality,

$$\sum_{i=1}^n |x_i|^2 \leq \sum_{i=1}^n |x_i| \sum_{i=1}^n |x_i| = \|x\|_1 \|x\|_1 = \|x\|_1^2$$

Since $\|x\|_2^2 \leq \|x\|_1^2$, $\|x\|_2 \leq \|x\|_1$

Then consider $\|x_1\| = \sum_{i=1}^n |x_i| = \sum_{i=1}^n |u_i v_i|$ where $u_i = 1 \ \forall i$ and $v_i = x_i$.

By Cauchy-Schwarz Inequality, $\sum_{i=1}^n |u_i v_i| \leq \left(\sum_{i=1}^n |u_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^n |v_i|^2\right)^{\frac{1}{2}}$,

so $\|x\|_1 \leq \sqrt{n}\|v\|_2 \Rightarrow \|x\|_1 \leq \sqrt{n}\|x\|_2$

Therefore $\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2$

Consider the case in which $\lambda = \max_i |x_i|$ is the only nonzero element of x. In this case $\|x\|_2 = \left(\sum_{i=1}^n |x_i|^2\right)^{\frac{1}{2}} = (\lambda^2)^{\frac{1}{2}} = \lambda$.

Even in the case that $\lambda = 0$, this implies $\lambda \leq \|x\|_2$. This could be considered the smallest case of $\|x\|_2$, thus $\max_i |x_i| \leq \|x\|_2$.

So $\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2$.

Similarly, consider the case in which every element is the same. Logically, this means each element can be called the maximum, λ.

Thus $\sqrt{n}\|x\|_2 = \sqrt{n}\left(\sum_{i=1}^n |x_i|^2\right)^{\frac{1}{2}} = \sqrt{n}\left(\sum_i^n (\lambda)^2\right)^{\frac{1}{2}} = \sqrt{n}(n\lambda^2)^{\frac{1}{2}} = \sqrt{n}(\sqrt{n}\lambda)$

$= n\lambda = n\left(\max_i |x_i|\right)$

Thus $\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2 \leq n\|x\|_\infty$

4. Let $A, B \in \mathbb{R}^{n \times n}$ and $C = AB$. Prove that the Frobenius norm is mutually consistent, i.e. $\|AB\|_F \leq \|A\|_F \|B\|_F$

Proof:

$\|AB\|_F = \|C\|_F = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} |c_{ij}|^2\right)^{\frac{1}{2}}$. Using the formula for matrix multiplication, $c_{ij} = \sum_{k} a_{ik}b_{kj}$, $\|C\|_F$ may be rewritten as

$$= \left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\sum_{k=1}^{n} a_{ik}b_{kj}\right|^2\right)^{\frac{1}{2}}.$$ By Cauchy Inequality for matrices,

$$\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\sum_{k=1}^{n} a_{ik}b_{kj}\right|^2\right)^{\frac{1}{2}} \leq \left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\left(\sum_{k=1}^{n} a_{ik}^2\right)\left(\sum_{k=1}^{n} b_{kj}^2\right)\right]\right)^{\frac{1}{2}}$$

$$= \left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sum_{k=1}^{n}|a_{ik}|^2\right)\cdot\left(\sum_{k=1}^{n}|b_{kj}|^2\right)\right)^{\frac{1}{2}} = \left(\sum_{i=1}^{n}\sum_{k=1}^{n}|a_{ik}|^2\sum_{j=1}^{n}\sum_{k=1}^{n}|b_{kj}|^2\right)^{\frac{1}{2}}$$

$$= \left(\sum_{i=1}^{n}\sum_{k=1}^{n}|a_{ik}|^2\right)^{\frac{1}{2}}\left(\sum_{j=1}^{n}\sum_{k=1}^{n}|b_{kj}|^2\right)^{\frac{1}{2}} = \|A\|_F\|B\|_F$$

Thus $\|C\|_F \leq \|A\|_F\|B\|_F$, meaning the Frobenius norm is mutually consistent.

5. Prove that A has rank one iff $A = uv^T$ for some nonzero vectors $u, v \in \mathbb{R}^n$.

Proof: Considering the rank-nullity theorem for 2 matrices

$A, B \in \mathbb{R}^{n\times n}$, $\underset{\text{dim(Col(A))}}{\text{rank}}(A) + \underset{\text{dim(Null(A))}}{\text{nullity}}(A) = n$ and $\text{rank}(B) + \text{nullity}(B) = n$.

Since $\text{Col}(AB) \subseteq \text{Col}(A)$, $\text{rank}(AB) \le \text{rank}(A)$. Similarly, since $\text{Null}(B) \subseteq \text{Null}(AB)$, $\text{nullity}(B) \le \text{nullity}(AB)$. B and AB have the same number of columns, so by the rank-nullity theorem,

$$\text{rank}(AB) + \text{nullity}(AB) = \text{rank}(B) + \text{nullity}(B) = n.$$

Since $\text{nullity}(B) \le \text{nullity}(AB)$, it follows that $\text{rank}(AB) \le \text{rank}(B)$.

$\text{rank}(AB) \le \text{rank}(A)$ and $\text{rank}(AB) \le \text{rank}(B)$ may be simplified to $\text{rank}(AB) \le \min(\text{rank}(A), \text{rank}(B))$ $\forall\ A, B \in \mathbb{R}^{n\times n}$.

Now considering $\text{rank}(A) = \text{rank}(uv^T)$ and the fact that $u \in \mathbb{R}^{n\times 1} \setminus \{0\}$

$\text{rank}(uv^T) \le \min(\text{rank}(u), \text{rank}(v^T)) \le 1$. Since both u and v are nonzero vectors, they must also both have at least rank 1. Thus $1 \le \text{rank}(uv^T) \le 1$. This is a strict requirement, since otherwise, $0 \le \text{rank}(uv^T) \le 1$.

So $\text{rank}(uv^T) = \text{rank}(A) = 1$

6. Show that if $A \in \mathbb{R}^{n \times n}$ is strictly upper triangular, then $A^n = 0$.

For $a_{ij} \in A$, $a_{ij} = 0$ for $i \geq j$, or equivalently,
$a_{ij} = 0$ for $i > j-1$.
Assume by induction that $a_{ij}^{k-1} = 0$ for $i > j-(k-1)$.
Then, $a_{ij}^{k} = 0$ for $i > j-k$ such that

$a_{ij}^k = (AA^{k-1})_{ij} = \sum_{m=1}^{n} a_{im} a_{mj}^{k-1}$. Splitting this into two series around i gives $\sum_{m=1}^{i} a_{im} a_{mj}^{k-1} + \sum_{m=i+1}^{n} a_{im} a_{mj}^{k-1}$.

For a_{im} in the first series, $i > m-1$, satisfying $i > j-k$, so $a_{im} = 0 \quad \forall m \leq i$. Similarly, a_{mj}^{k-1} in the second series $m \geq i+1 > (j-k)+1 \Rightarrow m > j-(k+1)$, so $a_{mj}^{k-1} = 0$ by the inductive hypothesis. So,

$$a_{ij}^k = \sum_{m=1}^{i} 0 \cdot a_{mj}^{k-1} + \sum_{m=i+1}^{n} a_{im} \cdot 0 = 0 \quad \forall i, j.$$

Therefore we can say $A^n = 0$.

```
import numpy as np
from time import perf_counter_ns

def getTridiagonalInverse(subD, mainD, supD):
    # great speedup - still room for improvement - play with indices later to
avoid separate loops
    n = len(mainD)
    theta = np.ones(n+1)
    psi = np.ones(n+1)
    theta[1] = mainD[0]
    psi[-2] = mainD[-1]
    for i in range(2,n+1):
        theta[i] = mainD[i-1]*theta[i-1] - supD[i-2]*subD[i-2]*theta[i-2]
    for i in range(n-2,-1,-1):
        psi[i] = mainD[i]*psi[i+1] - supD[i]*subD[i]*psi[i+2]

    Ainv = np.zeros((n,n))
    for i in range(n):
        for j in range(n):
            if i < j:
                Ainv[i,j] = (((-
1)**(i+j))*np.prod(supD[i:j])*theta[i]*psi[j+1])/theta[n]
            elif i > j:
                Ainv[i,j] = (((-
1)**(i+j))*np.prod(subD[j:i])*theta[j]*psi[i+1])/theta[n]
            else:
                Ainv[i,j] = (theta[i]*psi[j+1])/theta[n]
    return Ainv

def testInverse(A, inv):
    if not (np.allclose(np.linalg.inv(A), inv)):
        raise Exception(f'Something went wrong with new matrix inverse method')

def getTriDiagChol(subD, mainD, supD):
    n = len(mainD)
    A = np.zeros((n,n))
    diags = np.multiply(subD, supD)
    np.fill_diagonal(A, np.square(mainD))
    # works because fill_diagonal is in-place
    np.fill_diagonal(A[:(n-1), 1:], diags)
    np.fill_diagonal(A[1:, :(n-1)], diags)
    return A

def getMatrix(subD, mainD, supD):
    n = len(mainD)
```

```python
    A = np.zeros((n,n))
    np.fill_diagonal(A, mainD)
    # works because fill_diagonal is in-place
    np.fill_diagonal(A[:(n-1), 1:], supD)
    np.fill_diagonal(A[1:, :(n-1)], subD)
    return A

if __name__ == "__main__":
    subDiag = -1*np.ones(9)
    supDiag = -1*np.ones(9)
    mainDiag = 2*np.ones(10)
    B = getMatrix(subDiag, mainDiag, supDiag)

    new_inv = getTridiagonalInverse(subDiag, mainDiag, supDiag)
    testInverse(B, new_inv)
    cholFactor = getTriDiagChol(subDiag, mainDiag, supDiag)

    # a. Find condition number κ2(B)
    A_norm = np.sqrt(max(np.linalg.eigvals(cholFactor)))
    Ainv_norm = np.sqrt(max(np.linalg.eigvals(np.matmul(new_inv.T,new_inv))))
    condNumber = A_norm*Ainv_norm
    print(f'condition number κ2(B): {condNumber}')
    # b. Find the smallest (λmin) and the largest eigenvalues (λmax) of B to
        #compute the ratio λmax/λmin.
    eigenvals = np.linalg.eigvals(B)
    eigenvalRatio = max(eigenvals)/min(eigenvals)
    print(f'ratio λmax/λmin: {eigenvalRatio}')
    # c. compare the above results
    print('conclusion: ratio > condition number')
```

Output:

condition number κ2(B): 30.030550668916735
ratio λmax/λmin: 48.37415007870826
conclusion: ratio > condition number