APPLIED LINEAR ALGEBRA HOMEWORK 4

TOMMENIX YU
STAT 31430
DUE FRIDAY, NOV. 12, 3PM

Exercise 1. *Given two vectors $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ in $\mathbb{R}^n$ let $A = (a_{ij})$ be the matrix with entries given by*

$$a_{ij} = x_i y_j, \quad 1 \le i, j \le n.$$

(Note: This is the "rank 1" matrix $A = xy^T$, where we regard x and y as column vectors.)

(a) Show that for all $v \in \mathbb{R}^n$, $Av = \langle y, v\rangle x$, where $\langle y, v\rangle = y \cdot v$ is the usual Euclidean inner product. (This can be quick, but please explain your argument in complete sentences, and try to explain every step you make. If something seems too simple to explain, think about what definitions you are using.)

(b) For $p \le 1$, let $||A||_p$ denote the matrix norm subordinate to the vector norm $x \mapsto ||x||_p$. Determine conditions on x and y which characterize when

$$||A||_p = 1 \quad \textit{holds for all} \quad 1 \le p \le \infty.$$

Proof.

(a):

$$Av = xy^T v = x(y^T v) = x\langle y, v\rangle = \langle y, v\rangle x$$

where the equality is just because inner product yields a scalar.

(b): (discussed with Jason)

By definition of induced norms (subordinate norms) we have

$$||A||_p = \sup_v \frac{||Av||_p}{||v||_p} = \sup_v \frac{||\langle y, v\rangle x||_p}{||v||_p} = ||x||_p \sup_v \frac{\langle y, v\rangle}{||v||_p} = ||x||_p ||y||_{p'}$$

where the last equation is proved in HW3, Q6.

We can say further by Hölder's inequality that $\langle x, y\rangle \le ||x||_p ||y||_{p'} = 1$, given the condition $||A||_p = 1$.

Let condition (A) denote:

$$||x||_p ||y||_{p'} = 1, \quad \forall p \ge 1$$

and condition (B) denote the following:

$$x = rv \quad \text{and} \quad y = \frac{v}{r}$$

for $r > 0$ and any $v \in \mathbb{R}^n$ that satisfies

$$|v_{\sigma(1)}| = \cdots = |v_{\sigma(k)}| = \sqrt{\frac{1}{k}}, v_{\sigma_1(k+1)} = \cdots = v_{\sigma(n)} = 0$$

for $k \in \{1, \ldots, n\}$, $\sigma \in S_n$.

I will now prove that

$$(A) \iff (B).$$

$(A) \Rightarrow (B)$:

Since $||x||_p||y||_{p'} = 1$ holds for all p, and thus we know

$$\left(\sum_{i=1}^n |x_i|^2\right)^{1/2}\left(\sum_{i=1}^n |y_i|^2\right)^{1/2} = 1 \Rightarrow \left(\sum_{i=1}^n |x_i|^2\right)\left(\sum_{i=1}^n |y_i|^2\right) = 1$$

Now, on the one hand we have by 5(a), HW3

$$||x||_2^2||y||_2^2 \le ||x||_1||x||_\infty||y||_1||y||_\infty$$

and on the other hand

$$||x||_2^2||y||_2^2 = 1 = ||x||_1||x||_\infty||y||_1||y||_\infty$$

so the inequality $||x||_2^2 \le ||x||_1||x||_\infty$ takes equality, which means (from step in that homework)

$$\sum_{i=1}^n x_i^2 = \sum_{i=1}^n \left(|x_i| \cdot \max_{1\le j\le n} |x_j|\right)$$

i.e. all non-zero term in x has the same length. But since $x \in \mathbb{R}^n$ it means $x_{\sigma_1(1)} = \cdots = x_{\sigma_1(k_1)}, x_{\sigma_1(k_1+1)} = \cdots = x_{\sigma_1(n)} = 0$ for some permutation σ of $\{1, \ldots, n\}$ and some $k \in \{1, \ldots, n\}$. The same for y with k_2 non-zero terms and permutation σ_2.

Moreover, $||x||_2||y||_2 = 1$, which means we can assume $||x||_2 = r$ and $||y||_2 = \frac{1}{r}$.

We have for now:

$$|x_{\sigma_1(1)}| = \cdots = |x_{\sigma_1(k_1)}| = \sqrt{\frac{r}{k_1}}, x_{\sigma_1(k_1+1)} = \cdots = x_{\sigma_1(n)} = 0$$

and

$$|y_{\sigma_2(1)}| = \cdots = |y_{\sigma_2(k_2)}| = \sqrt{\frac{1}{rk_2}}, y_{\sigma_2(k_2+1)} = \cdots = y_{\sigma_2(n)} = 0$$

for $r > 0$, $k_1, k_2 \in \{1, \ldots, n\}$, $\sigma_1, \sigma_2 \in S_n$.

We now prove that $k_1 = k_2$ and $\sigma_1 \sim \sigma_2$, by which I mean that there exists some (not unique) σ such that it works for both x and y, and the negative/positive signs are chosen correspondingly, i.e. $x_i y_i \geq 0$.

To begin with, we use the above defined k_1 and k_2. Then since we know

$$||x||_1||y||_\infty = 1 = ||x||_\infty||y||_1$$

we get

$$\sum_{i=1}^{n} |x_i| = \frac{1}{|y_{\sigma_2(1)}|} \quad \text{and} \quad \sum_{i=1}^{n} |y_i| = \frac{1}{|x_{\sigma_1(1)}|}$$

but we also know that

$$\sum_{i=1}^{n} |x_i| = k_1|x_{\sigma_1(1)}| \quad \text{and} \quad \sum_{i=1}^{n} |y_i| = k_2|y_{\sigma_2(1)}|$$

thus

$$k_1 = \frac{1}{|x_{\sigma_1(1)}||y_{\sigma_2(1)}|} = k_2.$$

Now we set out to show the last bit, that the two permutations are, up to unimportant permutations, the same, and that the negative signs are co-appearing.

Since $A = xy^T$, $\text{rank}(A) = 1$ and thus there exists an eigenvalue with corresponding eigenvector u. Writing it out we have

$$Au = \langle y, u\rangle x = cu$$

which means that x is a multiple of u, hence also an eigenvector, so

$$Ax = \langle x, y\rangle x.$$

Now, let $\tilde{x} = \dfrac{x}{||x||_2}$, which is the normalized vector of x, we get

$$A\tilde{x} = \langle x, y\rangle \frac{x}{||x||_2}$$

but $||A||_2 = 1$ means that $||A\tilde{x}||_2 = ||\tilde{x}||_2 = 1$, we take 2 norm on both sides to get

$$1 = ||A\tilde{x}||_2 = \left\|\langle x, y\rangle \frac{x}{||x||_2}\right\|_2 = \langle x, y\rangle \frac{||x||_2}{||x||_2} = \langle x, y\rangle.$$

This means that we attain the equality condition of Holder's inequality. We can allude to that to get our result, but I'll just directly compute here.

Since we have

$$1 = \langle x, y\rangle = \sum_{i=1}^{n} x_i y_i \leq \sum_{i=1}^{n} |x_i||y_i| \leq k_1|x_{\sigma_1(1)}||y_{\sigma_2(1)}| = k_1\sqrt{\frac{r}{k_1}}\sqrt{\frac{1}{rk_1}} = 1 \tag{1}$$

where the second inequality is because both x and y has exactly k_1 non-zero terms (remember we've shown $k_1 = k_2$).

Since the left and right hand side are the same in (1), both inequalities attains equality.

For the first inequality to attain inequality, we need $x_i y_i = |x_i||y_i|$, and thus it cannot be the case where that one of x_i and y_i is negative and one is positive. Also, since we are in $\mathbb{R}^n$, we get $\forall i$,

$$\mathrm{sign}(x_i) = \mathrm{sign}(y_i)$$

(if we say $\mathrm{sign}(0) = \mathrm{sign}(0)$).

For the second inequality, it being an equality means that the index of all non-zero entries of x and y are the same. This means that there exist $\sigma \in S_n$ such that

$$x_{\sigma(i)} = 0 \iff y_{\sigma(i)} = 0.$$

Thus, we have come to the conclusion that

$$|x_{\sigma(1)}| = \cdots = |x_{\sigma(k)}| = \sqrt{\frac{r}{k}}, x_{\sigma_1(k+1)} = \cdots = x_{\sigma(n)} = 0$$

and

$$|y_{\sigma(1)}| = \cdots = |y_{\sigma(k)}| = \sqrt{\frac{1}{rk}}, y_{\sigma(k+1)} = \cdots = y_{\sigma(n)} = 0$$

for $r > 0$, $k \in \{1, \dots, n\}$, $\sigma \in S_n$, and that $\mathrm{sign}(x_i) = \mathrm{sign}(y_i)$.

But we immediately note this can be written in a much better form:

$$x = rv \quad \text{and} \quad y = \frac{v}{r}$$

for $r > 0$ and any $v \in \mathbb{R}^n$ that satisfies

$$|v_{\sigma(1)}| = \cdots = |v_{\sigma(k)}| = \sqrt{\frac{1}{k}}, v_{\sigma_1(k+1)} = \cdots = v_{\sigma(n)} = 0$$

for $k \in \{1, \dots, n\}$, $\sigma \in S_n$.

Hence, $(A) \Rightarrow (B)$.

$(B) \Rightarrow (A)$:

this is an easier direction. By computation we have

$$||x||_p ||y||_{p'} = r \cdot \frac{1}{r} ||v||_p ||v||_{p'}$$

so we only need to show that

$$||v||_p ||v||_{p'} = 1.$$

But we have

$$||v||_p = \left(\sum_{|v_i|^p} \right)^{\frac{1}{p}} = \left(k \cdot \left(\sqrt{\frac{1}{k}} \right)^p \right)^{\frac{1}{p}} = k^{\frac{1}{p} - \frac{1}{2}}$$

and similarily

$$||v||_{p'} = k^{\frac{1}{p'}-\frac{1}{2}}$$

which means

$$||v||_p||v||_{p'} = k^{(\frac{1}{p}-\frac{1}{2}+\frac{1}{p'}-\frac{1}{2})} = k^{(1-1)} = 1$$

which is exactly (A).

□

Exercise 2.

Fix $n \geq 2$ and let $||\cdot||$ denote a (sub-multiplicative) matrix norm on $\mathcal{M}_n(\mathbb{C})$.

(a) Show that for all $A \in \mathcal{M}_n(\mathbb{C})$ and all $k \geq 1$,

$$\rho(A) \leq ||A^k||^{1/k}.$$

(b) Show that for all $A \in \mathcal{M}_n(\mathbb{C})$ and $\varepsilon > 0$, there exists $K \geq 1$ such that for all $k \geq K$,

$$||A^k||^{1/k} \leq \rho(A) + \varepsilon.$$

(c) Conclude that for all $A \in \mathcal{M}_n(\mathbb{C})$ one has

$$\lim_{k\to\infty} ||A^k||^{1/k} = \rho(A).$$

Proof.

(a):

For $k \in \mathbb{Z}^*$, this is easy:

Let $\rho(A) = |\lambda|$ where λ is an eigenvalue of A, if x is the corresponding eigenvector then there exists y such that $xy^* \neq 0$. Then

$$A^k xy^* = \lambda^k xy^*$$

which implies

$$|\lambda^k|||xy^*|| = ||A^k y^*|| \leq ||A^k||||xy^*|| \quad \Rightarrow \quad |\lambda|^k \leq ||A^k||$$

which then means $\rho(A) \leq ||A^k||^{1/k}$.

(b): I follow the hint to do this.

Let $A_\varepsilon = \dfrac{1}{\rho(A)+\varepsilon}A$. Then, since we've just multiplicated A by a scalar, the eigenvectors of A_ε are the same as A, and eigenvalues are scaled by $\dfrac{1}{\rho(A)+\varepsilon}$. But then

$$\rho(A_\varepsilon) = \frac{\rho(A)}{\rho(A)+\varepsilon} < 1$$

which means $\forall x$, $A_\varepsilon^k x \to 0$ as $k \to \infty$, and further $||A^k|| \to 0$ as all norms are equivalent (to a subordinate norm, say). Then, $\exists K$ such that for all $k > K$,

$$||A_\varepsilon^k|| < 1 \quad \Rightarrow \quad ||A^k|| \leq (\rho(A)+\varepsilon)^k \quad \Rightarrow \quad ||A^k||^{1/k} \leq \rho(A)+\varepsilon.$$

(c): By above results, $\forall \varepsilon$,

$$\rho(A) \leq ||A^k||^{1/k} \leq \rho(A)+\varepsilon$$

for large enough k. Thus, taking $k \to \infty$ on both sides we have $\forall \varepsilon > 0$,

$$\rho(A) \leq \lim_{k\to\infty} ||A^k||^{1/k} \leq \rho(A)+\varepsilon$$

But since ε is arbitrary, the middle term cannot be larger than $\rho(A)$ (otherwise can find smaller ε than the distance since we are in $\mathbb{R}$), so

$$\lim_{k\to\infty} ||A^k||^{1/k} = \rho(A).$$

$\square$

Exercise 3. *Fix $n \leq 2$. For $A \leq \mathcal{M}_n(\mathbb{C})$ define*

$$r(A) := \sup\{|\langle Ax, x\rangle| : ||x|| = 1\}$$

as the numerical radius of A.

(a) Let $A \mapsto ||A||_2$ denote the matrix norm subordinate to the usual Euclidean vector norm $x \mapsto ||x||_2$. Show that for all $A \in \mathcal{M}_n(\mathbb{C})$,

$$\rho(A) \leq r(A) \leq ||A||_2 \leq 2r(A).$$

(b) Show that for any $A \in \mathcal{M}_n(\mathbb{C})$ and any unitary $U \in \mathcal{M}_n(\mathbb{C})$, one has

$$r(UAU^*) = r(A).$$

(c) Show that if $A \in \mathcal{M}_n(\mathbb{C})$ is normal then $r(A) = ||A||_2$, and that if $U \in \mathcal{M}_n(\mathbb{C})$ is unitary, then

$$r(U^{-1}) = r(U) = 1.$$

Proof.

(a):

$\rho(\mathbf{A}) \leq \mathbf{r}(\mathbf{A})$: by definition, $\rho(A)$ refers to the absolute value of the largest eigenvalue. But for any eigenvalue, one of it's corresponding eigenvector's normalized form x is on the unit ball and the length of Ax is $|\lambda|||x|| = |\lambda|$. By choosing the eigenvalue with the largest norm we get

$$\rho(A) = ||Ax_0||$$

where $||x_0|| = 1$ and is the corresponding eigenvector of $|\lambda_{max}|$.

Since $x_0 = \frac{1}{\lambda_{max}} Ax_0$, we know that

$$||Ax_0|| = \frac{||Ax_0||^2}{||Ax_0||} = \frac{\langle Ax_0, \lambda_{max} x_0\rangle}{|\lambda_{max}|} = \pm\langle Ax_0, x_0\rangle \leq |\langle Ax_0, x_0\rangle| \leq \sup_{||x||=1} |\langle Ax, x\rangle|$$

which means $\rho(A) \leq r(A)$.

$\mathbf{r}(\mathbf{A}) \leq ||\mathbf{A}||_\mathbf{2}$:

$$||A||_2 = \sup_{||x||=1} ||Ax|| \geq \sup_{||x||=1} \Big| ||Ax|| \cdot ||x|| \cdot \sin\theta \Big| =\geq \sup_{||x||=1} |\langle Ax, x\rangle| = r(A)$$

for $\theta =$ the angle between Ax and x.

$||\mathbf{A}||_\mathbf{2} \leq \mathbf{2r}(\mathbf{A})$: (I tried for a long time and didn't know how to use the hint... but I realized that part (c) can help. Discussed with Tim and Muyi Chen.)

Note that for any fixed x

$$|\langle Ax, x\rangle| = |\langle x, A^*x\rangle| = |\overline{\langle A^*x, x\rangle}| = |\langle A^*x, x\rangle|$$

and thus $|\langle Ax, x\rangle|$ and $|\langle A^*x, x\rangle|$ attain their sup at the same x, which means

$$\begin{aligned} 2r(A) &= \sup_{||x||=1} (|\langle Ax, x\rangle| + |\langle Ax, x\rangle|) \\ &= \sup_{||x||=1} (|\langle Ax, x\rangle| + |\langle A^*x, x\rangle|) \qquad (= r(A) + r(A^*)) \\ \text{(trig)} &\geq \sup_{||x||=1} |\langle Ax, x\rangle + \langle A^*x, x\rangle| \\ &= \sup_{||x||=1} |\langle (A + A^*)x, x\rangle| = r(A + A^*) \end{aligned}$$

and since $|\langle A^*x, x\rangle| = |-\langle A^*x, x\rangle|$ we can get by the same argument except that we use $-A^*$ instead of A:

$$2r(A) \geq r(A - A^*).$$

I will prove later in (c), independent of this, that for normal matrix, $||A||_2 = r(A)$. So we can just check $A + A^*$ and $A - A^*$ are normal. $A + A^*$ is normal because it's Hermitian, and

$$(A - A^*)(A - A^*)^* = (A - A^*)(A^* - A) = AA^* - A^2 - (A^*)^2 + A^*A$$

$$= (A^* - A)(A - A^*) = (A - A^*)^*(A - A^*)$$

so it's also normal. So

$$4r(A) \geq r(A + A^*) + r(A - A^*) = ||A + A^*||_2 + ||A - A^*||_2 \geq ||2A||_2 = 2||A||_2$$

where the second inequality is due to triangle inequality as $(A + A^*) + (A - A^*) = 2A$.

(b): Note that

$$\langle UAU^*x, x\rangle = \langle AU^*x, U^*x\rangle$$

so that

$$r(UAU^*) = \sup_{||x||=1} |\langle UAU^*x, x\rangle| = \sup_{||x||=1} |\langle AU^*x, U^*x\rangle|$$

yet since U unitary, it maps the unit circle to the unit circle, so there's a 1-1 correspondence between x and U^*x for $||x|| = 1$, which means that

$$\sup_{||x||=1} |\langle AU^*x, U^*x\rangle| = \sup_{||y||=1} |\langle Ay, y\rangle| = r(A)$$

and thus $r(UAU^*) = r(A)$.

(c): Assume A normal, then $A = UDU^*$ and

$$r(A) = r(UDU^*) = r(D) = \sup_{||x||=1} |\langle Dx, x\rangle| = \sup_{||x||=1} \left| \sum_{i=1}^{n} d_i x_i \right| = |\max_i d_i|1^2 + \sum_{j \neq i} d_j \cdot 0 = \max_i d_i$$

But $\max_i d_i = \rho(D) = \rho(A) = ||A||_2$, where the first equality is by definition, and the last equality is due to A normal.

For unitary U, we have U and U^{-1} are normal and thus

$$r(U) = ||U||_2 = 1 = ||U^{-1}||_2 = r(U^{-1})$$

which is what we want.

□

Exercise 4. *Let* $A \mapsto \exp(A) = \sum_0^\infty \frac{1}{i!}A^i$ *be the matrix exponential function as defined in class.*

(a) Show that for all $A \in \mathcal{M}_n(\mathbb{C})$ *and all invertible* $P \in \mathcal{M}_n(\mathbb{C})$*,*

$$\exp(PAP^{-1}) = P\exp(A)P^{-1}.$$

(b) Show that if $D = \mathrm{diag}(d1, \dots, dn)$ *with* $d_i \in \mathbb{C}$ *for* $i = 1, \dots, n$*, then*

$$\exp(D) = \mathrm{diag}(e^{d_1}, \dots, e^{d_n}).$$

Conclude that if $A = PDP^{-1}$ *is a diagonalization of* A*, then, writing* p_i *for the ith column of* P *(which is an eigenvector associated to the eigenvalue* d_i*; see, e.g. the additional problems on HW1) and expanding*

$$x = \sum_{i=1}^n \alpha_i p_i, \quad \alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{C}^n,$$

(which corresponds to setting $\alpha = P^{-1}x$*) we have*

$$\exp(A)x = \sum_{i=1}^n \exp(d_i)\alpha_i p_i.$$

(c) Show that for every $A \in \mathcal{M}_n(\mathbb{C})$*,*

$$\det(\exp(A)) = \exp(\mathrm{tr}(A)).$$

Proof.

(a): We have

$$\exp(PAP^{-1}) = \sum_{i=0}^\infty \frac{1}{i!}(PAP^{-1})^i = \sum_{i=0}^\infty \frac{1}{i!}PA^iP^{-1} = P\left(\sum_{i=0}^\infty \frac{1}{i!}A^i\right)P^{-1} = P\exp(A)P^{-1}.$$

(b): Since $D^i = \mathrm{diag}(d_1^i, \dots, d_n^i)$, so

$$e^D = \sum_{i=0}^\infty \frac{1}{i!}\mathrm{diag}(d_1^i, \dots, d_n^i) = \mathrm{diag}\left(\sum_{i=0}^\infty \frac{1}{i!}d_1^i, \dots, \sum_{i=0}^\infty \frac{1}{i!}d_n^i\right) = \mathrm{diag}(e^{d_1}, \dots, e^{d_n}).$$

And thus, with notations in problem we can get

$$e^Ax = Pe^DP^{-1}\sum_{i=1}^n \alpha_i p_i = \sum_{i=1}^n \alpha_i Pe^D(P^{-1}p_i) = \sum_{i=1}^n \alpha_i Pe^D e_i = \sum_{i=1}^n \alpha_i P(e^{d_i}e_i) = \sum_{i=1}^n \alpha_i e^{d_i}p_i$$

which is what we want.

(c):

Lemma 0.1. *The trace of a matrix is the sum of its eigenvalues counted with multiplicity. The determinant is the product of all eigenvalues.*

Proof. (Lemma 0.1)

Trace:

The $n-1$ degree term in P_A has coefficient $(-1)^n(\text{tr}(A))$ since in the computation of determinant, in order to get λ^{n-1}, we have to choose $n-1$ element on the diagonal, but the last multiplier can only be the constant part of the term left on the diagonal.

But on the other hand, $P_A = (-1)^n(\lambda - \lambda_1)\cdots(\lambda - \lambda_n)$, which has $n-1$ degree term's coefficient as the sum of the eigenvalues. Combined, this means that the trace is the sum of the eigenvalues. Thus we're done.

Determinant:

This is similar, note that P_A has 0 degree term equals to the summand where no λ is chosen in the definition of determinant, the 0 degree term is the determinant of the matrix.

By from $P_A = (-1)^n(\lambda-\lambda_1)\cdots(\lambda-\lambda_n)$ this expression it is also clear that the 0 degree term is nothing but the product of all eigenvalues, where the negative signs cancel out perfectly. $\square$

Now, by direct computing we have:

$$\det(e^A) = \det(Pe^T P^{-1})$$

where $A = PTP^{-1}$ is the Schur factorization, and T is an upper triangular matrix with its diagonal term being eigenvalues of A.

Now, note that for any upper triangular matrix U, UU is still upper triangular and has $UU_{ii} = u_{ii}^2$ where u_{ii} is the ith diagonal term of U. This is because

$$UU_{ii} = \sum_{j=1}^{i-1} u_{ji}\cdot 0 + u_{ii}^2 + \sum_{j=i+1}^{n} u_{ij}\cdot 0 = u_{ii}^2.$$

Given this, we know that the terms on the diagonal of e^T is going to be $e^{t_{ii}}$ for some corresponding i, because matrix addition is term by term, and the expansion is the same expansion for e^x. There fore we have

$$\det(e^A) = \det(Pe^T P^{-1}) = 1\cdot\det(e^T)\cdot 1 = \prod_{i=1}^{n} e^{t_{ii}} = \prod_{i=1}^{n} e^{\lambda_i} = e^{\sum_{i=1}^n \lambda_i} = \exp(\text{tr}(A))$$

where the last equality is due to lemma 0.1, trace part.

$\square$

Exercise 5.

One component of Strassen's 7-multiplication identity for multiplying $2x2$ matrices is an idea similar to an observation due to Gauss related to multiplication of complex numbers. Given $a, b, c, d \in \mathbb{R}$, show that if $z, w \in \mathbb{C}$ are given by

$$z = a + bi, \quad w = c + di,$$

then the product zw can be written in terms of $r = ac$, $s = bd$, and $t = (a + b)(c + d)$.

Let's now try to "rediscover" Strassen's identity. Consider the $2x2$ matrices

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad B = \begin{pmatrix} x & y \\ z & w \end{pmatrix}.$$

Set

$$\begin{cases} p = (b - d)(z + w), \\ q = (a + d)(x + w), \\ r = (a - c)(x + y), \\ s = (a + b)w, \\ t = a(y - w), \\ u = d(z - x), \\ v = (c + d)x. \end{cases}$$

Let C be the $2x2$ matrix given by the product $C = AB$, with entries $C = (c_{ij})_{i,j\in\{1,2\}}$. It turns out that the entries c_{ij} , can be written in terms of $p, \ldots, v$. This requires only seven multiplications, compared to the eight used in the "usual method."

Find c_{22} in terms of q, r, t, and v. (Although it is not needed for this part of the problem, for reference we identified c_{11} in class,and we also have $c_{12} = s + t$ and $c_{21} = u + v$.)

Note: Strassen's work has led to a lot of research on the algorithmic complexity of matrix mul- tiplication, continuing to the present day (see, e.g., D. Coppersmith and S. Winograd (1990), V. Vassilevska-Williams (2012), F. Le Gall (2014), J. Alman and V. Vassilevska-Williams (2018)).

Proof.

$$\begin{aligned} & q - r + t - v \\ =& (ax + aw + dx + dw) - (ax + ay - cx - cy) + (ay - aw) - (cx - dx) \\ =& dw + cy = c_{22}. \end{aligned}$$

□