Hwk #39:

This problem gives you the most celebrated use of Lagrange multipliers, but it requires some intrduction to appreciate it. (The calculations aren't bad at all.)

A famous task in linear algebra and matrix theory is to find *eigenvalues* of a given matrix. If A is a square matrix and you can find a *non-zero* vector v such that Av is actually a multiple of v, i.e., λv where λ is a number, then we call λ an eigenvalue of the matrix A (and v an eigenvector). For instance, the matrix $A = \begin{bmatrix} 3 & 2 \\ 3 & -4 \end{bmatrix}$ has 2 as an eigenvalue and $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ as a corresponding eigenvector, because

$$\begin{bmatrix} 3 & 2 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = 2 \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

It also has -3 as an eigenvalue with $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$ as an eigenvector. Of course multiples of eignevectors are again eigenvectors, e.g., if $Av = 2v$ then also $A(7v) = 2(7v)$. — In the example, there are only these two numbers $\lambda_1 = 2$ and $\lambda_2 = -3$ that are eigenvalues. If you try to find $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ solving $Av = \lambda v$ for any other λ you will only get the solution $v_1 = v_2 = 0$, i.e., only the zero vector. (Try it, just to gain familiarity with the notions.)

This problem is about eigenvalues of *symmetric* matrices. They play a role in studying definiteness of symmetric matrices. In physics, they are key concepts in describing rotating motions of rigid bodies. To every body, there is associated a symmetric 3×3 matrix called its 'tensor of inertia', whose eigenvectors point in the directions of such axes about which the body can rotate without wobbling (i.e., in a balanced way). The eigenvalues are called the moments of inertia about these axes.

To every symmetric $n \times n$ matrix A we associate the quadratic function $f(x) := x^T Ax$ where $x = [x_1, \ldots, x_n]^T \in \mathbb{R}^n$. We try to minimize or maximize $f(x)$ under the constraint $x^T x = 1$ (i.e., for x on the unit sphere).

(a) Write out $f(x)$ in components x_i for a 3×3 matrix A whose entries are called a_{ij}. Explain why a global maximum and a lobal minimum of $f(x)$ on the sphere are a-priori guaranteed to exist.

(b) Use the Lagrange multiplier method to set up equations satisfied by the x proociding a minimum or a maximum. (You may have written all these in components; but now make sure to rewrite the whole stuff again in matrix and vector form.) While you are not asked to actually solve for x (that would be very tedious, involving a cubic equation for λ), I ask you to express the value of f at the minimum and maximum in terms of the Lagrange multiplier. [Be aware that when finding the max vs the min, x and λ will typically refer to numerically different quantities in these two cases.]

You have just proved that every symmetric 3×3 matrix has (at least) two real eigenvalues. (Actually, if A is a multiple of the identity matrix, these two eigenvalues coincide.) And with jut a bit more writing, the same can be done for symmetric $n \times n$ matrices.

The method can be cranked up a bit, by throwing in further constraints, to prove that every symmetric $n \times n$ matrix has n real eigenvalues (some of which may coincide). This may well be among the most important pieces of insight in undergraduate mathematics, and it's a pity that it often falls between the cracks of separating Calc 3 and Linear Algebra into independent courses of the curriculum.

(c) Show, in a very brief calculation: If A is positive definite, then all its eigenvalues are positive. If A is positve semidefinte, then all of its eigenvalues are ≥ 0. FYI: The converse is also true; so indeed a symmetric matrix is positve definite (resp. emidefinite) IF AND ONLY IF all of its eigenvalues are positive (resp. non-negative). This statement is sponsored by the above proof (a), (b) and some extra dose of linear algebra. It is the launch pad for proving the Hurwitz and Gershgorin tests I gave you before.

Solution: (a) Using the symmetry of the matrix already, we get 6 pts

$$f(x_1, x_2, x_3) = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$
$$= a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3$$

Being a polynomial, this function is in particular continuous, and we constrain it to the sphere $x_1^2 + x_2^2 + x_3^2 = 1$, which is a closed and bounded set. Therefore a minimum and a maximum of f on the sphere exist.

(b) The Lagrange multiplier method says that at a constrained minimum, and at a constrained maximum, there exists a $\lambda \in \mathbb{R}$ such that $\nabla f = \lambda \nabla g$ (where $g(x) := x^T x - 1 = 0$ is the constraint), provided ∇g does not vanish on the set $g = 0$. We can easily see $\nabla g(x) = 2x$ (and this clearly does not vanish on the constraining set $x^T x = 1$). Let's calculate $\nabla f(x)$:

$$\nabla f(x) = \begin{bmatrix} \partial_1 f(x) \\ \partial_2 f(x) \\ \partial_3 f(x) \end{bmatrix} = \begin{bmatrix} 2a_{11}x_1 + 2a_{12}x_2 + 2a_{13}x_3 \\ 2a_{12}x_1 + 2a_{22}x_2 + 2a_{23}x_3 \\ 2a_{13}x_1 + 2a_{23}x_2 + 2a_{33}x_3 \end{bmatrix} = 2Ax$$

We conclude there exist $\lambda \in \mathbb{R}$ and a nonzero vector $x \in \mathbb{R}^3$ such that $2Ax = 2\lambda x$, hence λ i an eigenvalue and x is an eigenvector. (We know more about x than that it is non-zero; actually $\|x\| = 1$.)

Since $Ax = \lambda x$ at the extremum, we conclude $x^T Ax = \lambda x^T x = \lambda$, so the eigenvalue is actually the value of the constrianed maximum / minimum of f. Unless these two coincide (in which case $x^T Ax$ would have to be constant on the sphere, i.e., A would have to be a multiple of the unit matrix I), we indeed have found two different solutions λ (one for the min and one for the max).

(c) If A has an eigenvalue λ (with eigenvector $x \neq 0$), i.e., if $Ax = \lambda x$, then $x^T Ax = \lambda \|x\|^2$. For a positive definite matrix A, this expression must be positive for *all* vectors $x \neq 0$, in particular for the eigenvector x Hence $\lambda > 0$. Similarly we can argue for positive semidefinite.

Hwk #40:

Think of the task of finding the absolute maximum of $x^2 + \frac{1}{2}y^2 + y^4 - xy$ on the set S given by $(x-1)^2 + |y + y^3| \leq 5$. The purpose of this problem is NOT that you would actually do calculations to FIND the maximum (which would require numerical methods). Rather, in preparation for such a search. I want you to use the Hessian to conclude that the maximum exists and is on the BOUNDARY of the set S.

The message here is: While modest problems can already lead to prohibitively complicated calculations that may need numerical tools, simple analytic arguments may still be able significantly to reduce the amount of labor in a numerical search.

Solution: First we note that an absolute maximum exists, b/c we have a continuous expression on a bounded and closed set. We argue that any absolute maximum in this case cannt be in the interior because the Hessian is not negative semidefinite anywhere. (We do not attempt to solve the equations from vanishing of the gradient, even though we might of course have tried this, too. 4 pts

The Hessian is $\begin{bmatrix} 2 & -1 \\ -1 & 1+12y^2 \end{bmatrix}$. It is clearly positive definite everywhere by the Hurwitz test. So if a critical point were to be found in the interior of S, it would not be an absolute maximum (but a relative minimum at least; possibly an absolute minimum).

Hwk #41:

> A variation of the theme from the previous problem that is more 'real life', coming straight from partial differential equations. Suppose we have a bounded domain Ω in $\mathbb{R}^3$ and a function f that is continuous on the closure $\bar{\Omega}$ and C^2 in Ω. Suppose f satisfies $\partial^2 f(x,y,z)/\partial x^2 + \partial^2 f(x,y,z)/\partial y^2 + \partial^2 f(x,y,z)/\partial z^2 = 1$ in Ω, and $f = 0$ on the boundary. Show that then $f < 0$ in Ω. *Hint: write down the diagonal of the Hessian and show that necessary conditions for the negative definiteness are incompatible with the information given about f.* — Note: This kind of reasoning is known under the name of 'maximum principle'

Solution: Suppose there were some $x \in \Omega$, where $f(x) \geq 0$. The maximum of f on the closed set $\bar{\Omega}$ exists. It is obviously ≥ 0, because $f = 0$ on the boundary part of $\bar{\Omega}$. If this maximum is > 0, it must clearly be taken on in some point x_0 in the interior Ω. If the maximum however is 0 and f is nonnegative at some point in the interior, then we would again have a maximum in the interior. At such a point the Hessian would have to be negative semidefinite. This implies in particular that the diagonal elements of the Hessian would have to be ≤ 0. (Further conditions would be required from the off-diagonal entries of the Hessian, but we have not studied these and do not need or use them here.) 5 pts

However, the PDE $\partial^2 f(x,y,z)/\partial x^2 + \partial^2 f(x,y,z)/\partial y^2 + \partial^2 f(x,y,z)/\partial z^2 = 1$ says that the sum of the diagonal elements of the Hessian in *any* point inside Ω (in particular the maximum point x_0) has to be 1, hence cannot be the sum of numbers that are≤ 0.

This contradiction shows that the assumption $f(x) \geq 0$ for some $x \in \Omega$ must have been false.

Hwk #42:

> Let T be the set $\{(x,y) \mid 0 \leq x \leq \pi,\ 0 \leq y \leq x\}$. Draw a figure of this set. Then evaluate the integral $I := \int_T \sin x \, \sin y \, d(x,y)$ in two ways: as iterated integral in either order.
>
> *Hint: Make sure you get the limits of integration right. If any of your calculations leaves a dangling x or y in the result you sure haven't gotten the limits right. This alert applies to all MV integral problems.*

Solution: T is a triangle: ⊿. We can integrate over y first: for each fixed x, we note that y runs from 0 to x. These y integrals occur for x from 0 to π. So, 5 pts

$$I = \int_0^\pi \Big(\int_0^x \sin x \, \sin y \, dy \Big) \, dx = \int_0^\pi (1 - \cos x) \sin x \, dx = 2 \,.$$

Alternatively, we can integrate over x first. Then, for each fixed y, the x-integral extends from y to π. So we have

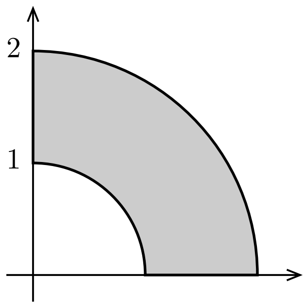
$$I = \int_0^\pi \Big(\int_y^\pi \sin x \, \sin y \, dx \Big) \, dy = \int_0^\pi (1 + \cos y) \sin y \, dy = 2 \, .$$

Hwk #43:

Let A be the set $\{(x, y) \mid 1 \leq x^2 + y^2 \leq 4, \; x, y \geq 0\}$. Draw a figure of this set. Then evaluate the integral $I := \int_A x^2 y \, d(x, y)$ in two ways: one version using cartesian coordinates, and one using polar coordinates.

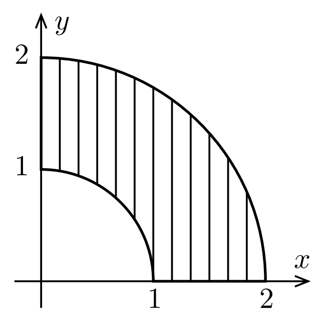
Using cartesian coordinates here is a bit dumb, admittedly. But I am asking that you do it anyways, to see the comparison with polar coordinates, and as a training to deal with the limits of integration correctly. Note that one order of integration in cartesian coordinates is easier to calculate than the other. Can you see which, and why?

Solution: 5 pts



The domain is a quarter of an annulus as depicted on the left. Integration in cartesian coordinates requires splitting the outer integral into two, because the limits of integration in the inner integral are given by a piecewise function. We'll carry out both orders of integral in cartesian coordinates, for illustration purposes. But using $\int \ldots dy$ as the inner integral is easier, because the antiderivative $y^2/2$ makes the square roots disappear from the outer integral.

First we consider the integration with y as the inner integral. If $0 \leq x \leq 1$, the y integral extends between the two circular arcs, namely from $y = \sqrt{1 - x^2}$ to $y = \sqrt{4 - x^2}$. If $1 \leq x \leq 2$, the y integral extends from $y = 0$ to $y = \sqrt{4 - x^2}$. So the outer integral needs to be split into $\int_0^1 \ldots dx + \int_1^2 \ldots dx$.



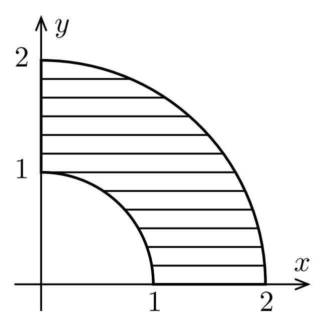
$$I = \int_0^1 \left(\int_{\sqrt{1-x^2}}^{\sqrt{4-x^2}} x^2 y \, dy \right) dx + \int_1^2 \left(\int_0^{\sqrt{4-x^2}} x^2 y \, dy \right) dx$$

This evaluates to

$$I = \int_0^1 x^2 \left[\frac{y^2}{2} \right]_{y=\sqrt{1-x^2}}^{y=\sqrt{4-x^2}} dx + \int_1^2 x^2 \left[\frac{y^2}{2} \right]_{y=0}^{y=\sqrt{4-x^2}} dx = \int_0^1 \frac{3}{2} x^2 \, dx + \int_1^2 \frac{1}{2} x^2 (4 - x^2) \, dx = \frac{1}{2} + \frac{14}{3} - \frac{31}{10}$$

Doing the fractions we get $I = \frac{31}{15}$.

Next we consider the integration with x as the inner integral. If $0 \leq y \leq 1$, the x integral extends between the two circular arcs, namely from $x = \sqrt{1 - y^2}$ to $x = \sqrt{4 - y^2}$. If $1 \leq y \leq 2$, the x integral extends from $x = 0$ to $x = \sqrt{4 - y^2}$. So the outer integral needs to be split into $\int_0^1 \ldots dy + \int_1^2 \ldots dy$.

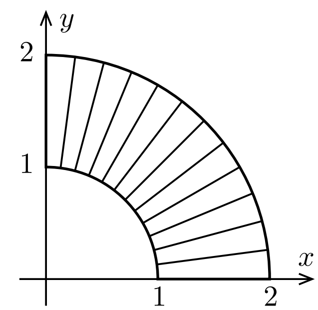


$$I = \int_0^1 \left(\int_{\sqrt{1-y^2}}^{\sqrt{4-y^2}} x^2 y \, dx \right) dy + \int_1^2 \left(\int_0^{\sqrt{4-y^2}} x^2 y \, dx \right) dy$$

This evaluates to

$$
\begin{aligned}
I &= \int_0^1 y\left[\frac{x^3}{3}\right]_{x=\sqrt{1-y^2}}^{x=\sqrt{4-y^2}} dy + \int_1^2 y\left[\frac{x^3}{3}\right]_{x=0}^{x=\sqrt{4-y^2}} dy \\
&= \int_0^1 \frac{y}{3}\left((4-y^2)^{3/2}-(1-y^2)^{3/2}\right) dy + \int_1^2 \frac{y}{3}(4-y^2)^{3/2} dy \\
&= -\frac{1}{15}\left[(4-y^2)^{5/2}-(1-y^2)^{5/2}\right]_0^1 - \frac{1}{15}\left[(4-y^2)^{5/2}\right]_1^2 = \frac{4^{5/2}}{15}-\frac{1}{15}=\frac{31}{15}
\end{aligned}
$$

Finally, we evaluate the integral in polar coordinates, using $x = r\cos\varphi$, $y = r\sin\varphi$ and $d(x,y) = r\,dr\,d\varphi$. DON'T FORGET THE r IN THE DIFFERENTIAL! The limits of integration are much easier now: $1 \le r \le 2$ and $0 \le \varphi \le \frac{\pi}{2}$. Order of integration is inessential: since the integrand is a product of an r function and a φ function, both single variable integrals can be evaluated independently of each other.



$$I = \int_0^{\pi/2}\int_1^2 (r^3\cos^2\varphi\,\sin\varphi)\,r\,dr\,d\varphi$$

We can move the trigs in front of the r-integral. Then the remaining r integral is a constant, not depending on φ, and it can therefore be moved out of the φ integral. Look carefully at this procedure: it will occur often, and at first glance it may look like a "product rule $\int fg = (\int f)(\int g)$". NOT SO! The essence is that the integrand is a product of functions each of which depends on a *different* variable, and that the limits of each integration do not depend on the other variables. It is ONLY then that you can use this trick. If you want to have a name for this rule (there doesn't seem to be an official name), call it the 'TENSOR PRODUCT RULE'.

$$I = \int_0^{\pi/2}\left(\cos^2\varphi\,\sin\varphi\int_1^2 r^4\,dr\right)d\varphi = \left(\int_1^2 r^4\,dr\right)\left(\int_0^{\pi/2}\cos^2\varphi\,\sin\varphi\,d\varphi\right) = \frac{31}{5}\cdot\frac{1}{3}[-\cos^3\varphi]_0^{\pi/2} = \frac{31}{15}$$

Hwk #44:

Let $\vec{u} = [2,1,3]^T$, $\vec{v} = [-1,0,4]^T$, $\vec{w} = [2,-1,-3]^T$. Calculate $\vec{v}\times\vec{w}$, $\vec{u}\times(\vec{v}\times\vec{w})$, $\vec{u}\times\vec{v}$, $(\vec{u}\times\vec{v})\times\vec{w}$.

With these same vectors from the previous problem, calculate $\vec{u}\cdot(\vec{v}\times\vec{w})$ and $\vec{w}\cdot(\vec{u}\times\vec{v})$.

Solution: 5 pts

$$\vec{v}\times\vec{w} = \begin{bmatrix}-1\\0\\4\end{bmatrix}\times\begin{bmatrix}2\\-1\\3\end{bmatrix} = \begin{bmatrix}0(-3)-(-1)4\\4\cdot2-(-1)(-3)\\(-1)(-1)-0\cdot2\end{bmatrix} = \begin{bmatrix}4\\5\\1\end{bmatrix}$$

$$\vec{u}\times(\vec{v}\times\vec{w}) = \begin{bmatrix}2\\1\\3\end{bmatrix}\times\begin{bmatrix}4\\5\\1\end{bmatrix} = \begin{bmatrix}1\cdot1-3\cdot5\\3\cdot4-2\cdot1\\2\cdot5-1\cdot4\end{bmatrix} = \begin{bmatrix}-14\\10\\6\end{bmatrix}$$

$$\vec{u}\times\vec{v} = \begin{bmatrix}2\\1\\3\end{bmatrix}\times\begin{bmatrix}-1\\0\\4\end{bmatrix} = \begin{bmatrix}1\cdot4-3\cdot0\\3(-1)-2\cdot4\\2\cdot0-1(-1)\end{bmatrix} = \begin{bmatrix}4\\-11\\1\end{bmatrix}$$

$$(\vec{u}\times\vec{v})\times\vec{w} = \begin{bmatrix} 4 \\ -11 \\ 1 \end{bmatrix} \times \begin{bmatrix} 2 \\ -1 \\ -3 \end{bmatrix} = \begin{bmatrix} (-11)(-3)-1(-1) \\ 1\cdot 2 - 4(-3) \\ 4(-1)-(-11)2 \end{bmatrix} = \begin{bmatrix} 34 \\ 14 \\ 18 \end{bmatrix}$$

(If you find it still difficult to safely remember the pattern of which to multiply with which, a quick check towards correctness of your calculation is to confirm that the dot product of the result with either factor is 0.)

Specifically our example has shown that $\vec{u}\times(\vec{v}\times\vec{w})$ *and* $(\vec{u}\times\vec{v})\times\vec{w}$ *are different!*

Now we calculate

$$\vec{u}\cdot(\vec{v}\times\vec{w}) = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 5 \\ 1 \end{bmatrix} = 2\cdot 4 + 1\cdot 5 + 3\cdot 1 = 16$$

$$\vec{w}\cdot(\vec{u}\times\vec{v}) = \begin{bmatrix} 2 \\ -1 \\ -3 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ -11 \\ 1 \end{bmatrix} = 2\cdot 4 + (-1)(-11) + (-3)\cdot 1 = 16$$

As they should, these two are equal.

Hwk #45:

Find the area of the triangle whose vertices are the points $A(1,1,3)$, $B(-2,3,0)$, $C(1,1,-2)$.

Solution: $\vec{AB} = [-3,\ 2,\ -3]^T$ and $\vec{AC} = [0,\ 0,\ -5]^T$. The area of the triangle is half the area of the parallelogram spanned by these two vectors, and this parallelogram area is the norm of the cross product of the two vectors. So the area is 3 pts

$$\frac{1}{2}\left\| \begin{bmatrix} 10 \\ 15 \\ 0 \end{bmatrix} \right\| = \frac{1}{2}\sqrt{325} = \frac{5}{2}\sqrt{13}$$

Note: You could have done the same kind of calculation (eg) with $\vec{BA}$ and $\vec{BC}$. Different route, but the result would have to be the same. Try it out if you wish.

Hwk #46:

Given the vectors $\vec{u} = [u_1, u_2, u_3]^T$ and $\vec{v} = [v_1, v_2, v_3]^T$ in space, I have defined their cross product $\vec{u}\times\vec{v}$ to be the vector $\vec{w} = [u_2v_3 - u_3v_2\,,\ u_3v_1 - u_1v_3\,,\ u_1v_2 - u_2v_1]^T$. Show that indeed, $\|\vec{w}\|^2 = \|\vec{u}\|^2\|\vec{v}\|^2(1-\cos^2\varphi)$, where φ is the angle between $\vec{u}$ and $\vec{v}$.

Solution: The right side is $\|\vec{u}\|^2\|\vec{v}\|^2 - (\vec{u}\cdot\vec{v})^2$. Now let's evaluate either side in components: 4 pts

$$\begin{aligned}\|\vec{w}\|^2 &= (u_2v_3 - u_3v_2)^2 + (u_3v_1 - u_1v_3)^2 + (u_1v_2 - u_2v_1)^2 = \\ &= u_2^2v_3^2 + u_3^2v_2^2 + u_3^2v_1^2 + u_1^2v_3^2 + u_1^2v_2^2 + u_2^2v_1^2 - 2u_2u_3v_2v_3 - 2u_1u_3v_1v_3 - 2u_1u_2v_1v_2\end{aligned}$$

$$\begin{aligned}\|\vec{u}\|^2\|\vec{v}\|^2 - (\vec{u}\cdot\vec{v})^2 &= (u_1^2+u_2^2+u_3^2)(v_1^2+v_2^2+v_3^2) - (u_1v_1+u_2v_2+u_3v_3)^2 \\ &= u_2^2v_3^2 + u_3^2v_2^2 + u_3^2v_1^2 + u_1^2v_3^2 + u_1^2v_2^2 + u_2^2v_1^2 - 2u_2u_3v_2v_3 - 2u_1u_3v_1v_3 - 2u_1u_2v_1v_2\end{aligned}$$