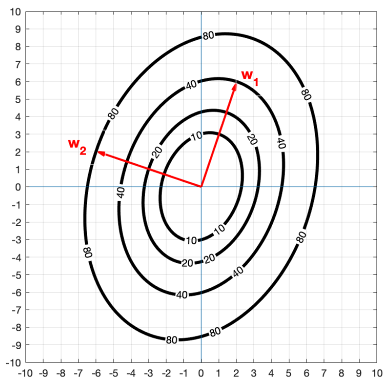
Solutions to Math 51 Practice problems for Quiz 9

1. (3 points) Each picture below is a contour plot of $q_M(x, y)$ for some symmetric matrix M; the contour labels indicate the values of $q_M(x, y)$. In each case, two eigenvectors $\mathbf{w}_1$ and $\mathbf{w}_2$ of M are also sketched, where

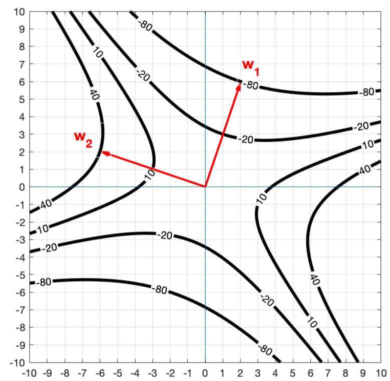
$$\|\mathbf{w}_1\| = \|\mathbf{w}_2\| = 2\sqrt{10}, \qquad M\mathbf{w}_1 = \lambda_1\mathbf{w}_1, \qquad M\mathbf{w}_2 = \lambda_2\mathbf{w}_2.$$

Determine $a = \frac{\lambda_2}{\lambda_1}$ for contour plot A, and $b = \frac{\lambda_2}{\lambda_1}$ for contour plot B.

(A)



(B)



Note that
For (A): $q_M(\mathbf{w}_1) = 40$ and $q_M(\mathbf{w}_2) = 80$.
For (B): $q_M(\mathbf{w}_1) = -80$ and $q_M(\mathbf{w}_2) = 40$.

(a) 0 (b) 1 (c) -1 (d) 2 (e) -2
(f) 3 (g) -3 (h) 4 (i) -4 (j) 1/2
(k) $-1/2$ (l) 1/3 (m) $-1/3$ (n) 1/4 (o) $-1/4$

For contour plot A, we note that the level curves of $q_M(x, y)$ are ellipses, so the two eigenvalues λ_1 and λ_2 of M must be either both positive or both negative, so the ratio $\dfrac{\lambda_2}{\lambda_1}$ must be positive.

By Equation 24.2.2 from the textbook, we have

$$q_M(c_1\mathbf{w}_1 + c_2\mathbf{w}_2) = \lambda_1 c_1^2\|\mathbf{w}_1\|^2 + \lambda_2 c_2^2\|\mathbf{w}_2\|^2.$$

Since $\|\mathbf{w}_1\| = \|\mathbf{w}_2\| = 2\sqrt{10}$,

$$q_M(c_1\mathbf{w}_1 + c_2\mathbf{w}_2) = (\lambda_1 c_1^2 + \lambda_2 c_2^2)(40).$$

When $c_1 = 1$ and $c_2 = 0$, we have

$$q_M(\mathbf{w}_1) = \lambda_1(40) = 40, \qquad \lambda_1 = 1.$$

When $c_1 = 0$ and $c_2 = 1$, we have

$$q_M(\mathbf{w}_2) = \lambda_2(40) = 80, \qquad \lambda_2 = 2.$$

Wee see that

$$\frac{\lambda_2}{\lambda_1} = 2.$$

The ellipse is longer in the direction of $\mathbf{w}_1$ corresponding to the eigenvalue with has smaller absolute value.

For contour plot B, we note that the level curves of $q_M(x, y)$ are hyperbolas, so the two eigenvalues λ_1 and λ_2 of M must have opposite signs, so the ratio $\dfrac{\lambda_2}{\lambda_1}$ must be negative.

By Equation 24.2.2 from the textbook, we have

$$q_M(c_1\mathbf{w}_1 + c_2\mathbf{w}_2) = \lambda_1 c_1^2\|\mathbf{w}_1\|^2 + \lambda_2 c_2^2\|\mathbf{w}_2\|^2.$$

Since $\|\mathbf{w}_1\| = \|\mathbf{w}_2\| = 2\sqrt{10}$,

$$q_M(c_1\mathbf{w}_1 + c_2\mathbf{w}_2) = (\lambda_1 c_1^2 + \lambda_2 c_2^2)(40).$$

When $c_1 = 1$ and $c_2 = 0$, we have

$$q_M(\mathbf{w}_1) = \lambda_1(40) = -80, \qquad \lambda_1 = -2.$$

When $c_1 = 0$ and $c_2 = 1$, we have

$$q_M(\mathbf{w}_2) = \lambda_2(40) = 40, \qquad \lambda_2 = 1.$$

Wee see that

$$\frac{\lambda_2}{\lambda_1} = -1/2.$$

The asymptotes of the hyperbolas are nearer $\mathbf{w}_2$ corresponding to the eigenvalue with the smaller absolute value.

2. (3 points) For which values b does the function

$$f(x, y) = 3x^2 - 2bxy + 12y^2$$

attain its global minimum value at *exactly one* point in $\mathbf{R}^2$?

(a) $b < 0$ (b) $b > 0$ (c) $|b| > 6$ (d) $-6 < b < 6$

The quadratic function f is the quadratic form associated with the symmetric matrix

$$A = \begin{bmatrix} 3 & -b \\ -b & 12 \end{bmatrix}.$$

This function has gradient $\nabla f = \begin{bmatrix} 6x - 2by \\ -2bx + 24y \end{bmatrix}$, which vanishes when $x = (b/3)y$ and $y = (b/12)x$; for most values of b, these two equations together force there to be a unique critical point $(x, y) = (0, 0)$. However, if $\frac{b}{3} = \frac{12}{b}$, which holds when $b^2 = 36$, or $b = \pm 6$, there will be infinitely many critical points lying along a line through the origin; we'll deal with this special case separately.

Regardless the value of b, the quadratic form f will have $(0, 0)$ as a critical point; and at this point, f attains the value 0. If the value of b is such that the quadratic form is *positive-definite*, then f will attain both a local, and a global, minimum *only* at $(0, 0)$. On the other hand, if f is negative or zero at some point away from the origin, then it is impossible for f to attain its global minimum value *only* at the origin; so we do not want f to be negative-definite, indefinite, or positive-semidefinite or negative-semidefinite.

The quadratic form is positive definite exactly when A's eigenvalues λ_1 and λ_2 are both positive. The eigenvalues of A are roots of the characteristic polynomial

$$P_A(\lambda) = \lambda^2 - 15\lambda + (36 - (-b)^2,$$

that is,

$$\lambda = \frac{15 \pm \sqrt{225 - 4(36 - b^2)}}{2} = \frac{15 \pm \sqrt{81 + 4b^2}}{2}.$$

In order for both of these eigenvalues to be positive, we must have $\sqrt{81 + 4b^2} < 15$. This is equivalent to requiring $81 + 4b^2 < 225$, equivalently $4b^2 < 144$, equivalently $b^2 < 36$. This is true when $\boxed{-6 < b < 6}$.

Alternatively, we find that $\text{tr(A)} = 3 + 12 = 15 > 0$, and $\det(A) = 3 \cdot 12 - (-b)^2 = 36 - b^2$. By Theorem 26.3.3, note that $\lambda_1 + \lambda_2 = 15 > 0$, so λ_1 and λ_2 will be both positive when $\det(A) > 0$, i.e. $36 - b^2 > 0$. This occurs when $|b| < 6$, or $\boxed{-6 < b < 6}$.

Meanwhile, we find that we can disregard the case $b = \pm 6$ (i.e., when f has infinitely many critical points, lying along the line $y = (b/12)x$), because the reasoning above shows that the eigenvalues of A are

$$\lambda = \frac{1}{2}\left(15 \pm \sqrt{81 + 4(36)}\right) = \frac{1}{2}\left(15 \pm \sqrt{225}\right) = 0, 15.$$

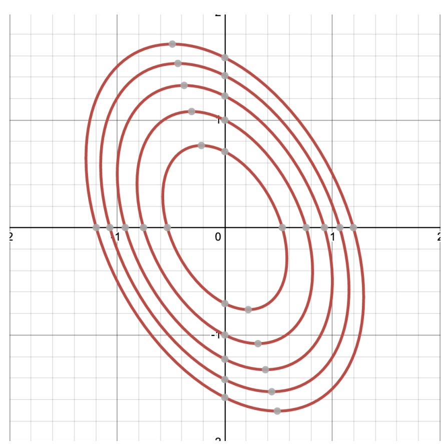
This means that f is positive-semidefinite but not positive-definite; in other words, $f \geq 0$ at all points in $\mathbf{R}^2$, but $f = 0$ at some points other than the origin. As we noted, such a function does *not* attain its global minimum value (i.e., 0) at exactly one point.

3. (4 points) Below are contour plots of **quadratic approximations at the origin** of four functions; each of the four functions has a critical point at $(0, 0)$. For each of the given Hessian matrices at the origin, select the picture representing the associated contour plot.

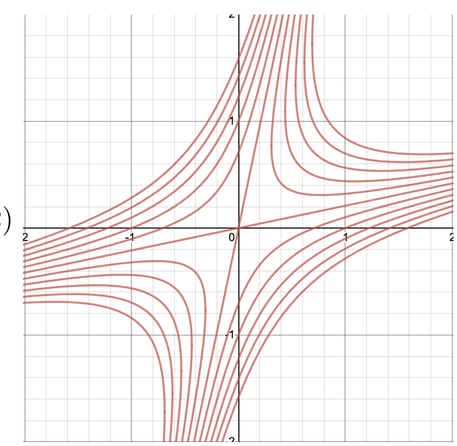
(a) $(\text{Hf})(0,0) = \begin{bmatrix} 14 & 4 \\ 4 & 8 \end{bmatrix}$

(b) $(\text{Hg})(0,0) = \begin{bmatrix} 4 & 10 \\ 10 & 4 \end{bmatrix}$

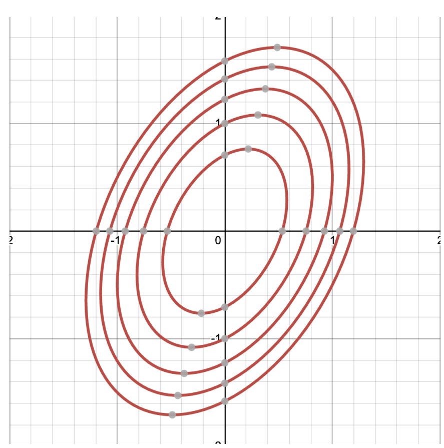
(1)



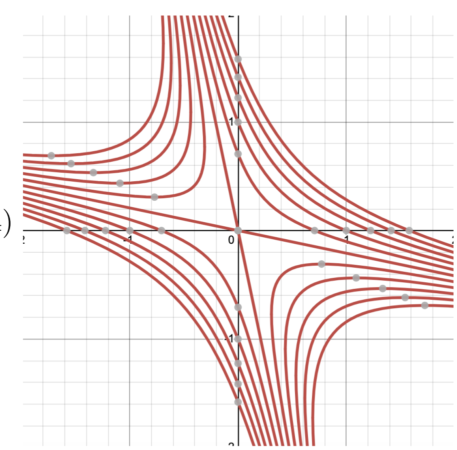
(2)



(3)



(4)



(A) Since f has a critical point at $(0,0)$, the quadratic approximation of f at the origin takes the form

$$\begin{aligned} f(h,k) &\approx f(0,0) + \frac{1}{2}\begin{bmatrix} h & k \end{bmatrix} ((\mathrm{Hf})(0,0)) \begin{bmatrix} h \\ k \end{bmatrix} \\ &= f(0,0) + \frac{1}{2}\begin{bmatrix} h & k \end{bmatrix} \begin{bmatrix} 14 & 4 \\ 4 & 8 \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix} \\ &= f(0,0) + 7h^2 + 4hk + 4k^2 \\ &= f(0,0) + q_A(h,k), \end{aligned}$$

where A is the matrix

$$A = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix}.$$

Level sets of f's quadratic approximation resemble (up to a re-labeling of levels) level sets of q_A. The matrix A has trace 11 and determinant $28 - 4 = 24$, so it has characteristic polynomial

$$\lambda^2 - 11\lambda + 24 = (\lambda - 8)(\lambda - 3),$$

which has as its roots $\lambda_1 = 8$ and $\lambda_2 = 3$ (which can also be found by the quadratic formula, if you didn't notice how it factors).

To sketch the contour plot, we need to work out perpendicular eigenvectors $\mathbf{w}_1$ and $\mathbf{w}_2$ for λ_1 and λ_2 respectively so as to write the quadratic form q_A in a more convenient reference frame. The lines for the eigenvalues are the null spaces of $A - 8\mathrm{I}_2$ and $A - 3\mathrm{I}_2$. We compute these matrices to be

$$A - 8\mathrm{I}_2 = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix}, \qquad A - 3\mathrm{I}_2 = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}.$$

The first has null space corresponding to the pair of equations $-x+2y = 0$ and $2x-4y = 0$, which are scalar multiples of each other (as they must be for a line of an eigenvalue): this is the line $y = (1/2)x$, so it is the span of the vector $\mathbf{w}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ (or any nonzero scalar multiple of this). The second of these has null space corresponding to the pair of equations $4x + 2y = 0$ and $2x + y = 0$, which are likewise scalar multiples of each other (as they must be for a line of an eigenvalue): this is the line $y = -2x$, so it is the span of the vector $\mathbf{w}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ (or any nonzero scalar multiple of this). The quadratic form q_A is given by the formula

$$q_A(c_1\mathbf{w}_1 + c_2\mathbf{w}_2) = \lambda_1(\mathbf{w}_1 \cdot \mathbf{w}_1)c_1^2 + \lambda_2(\mathbf{w}_2 \cdot \mathbf{w}_2)c_2^2 = 8(5)c_1^2 + 3(5)c_2^2.$$

This has level curves that are ellipses centered at $(0,0)$ with symmetry lines through $\mathbf{w}_1$ and $\mathbf{w}_2$ and stretched out more along the $\mathbf{w}_2$-line than along the $\mathbf{w}_1$-line (since $40c_1^2 + 15c_2^2 = c$ crosses the c_1-axis at $\pm\sqrt{c/40}$ and the c_2-axis at $\pm\sqrt{c/15}$, so the ratio of the length along the $\mathbf{w}_2$-line to the length along the $\mathbf{w}_1$-line is $\sqrt{c/15}/\sqrt{c/40} = \sqrt{40/15} > 1$). The ellipse is longer in the direction of $\mathbf{w}_2$ since $|\lambda_2| < |\lambda_1|$.

Thus, $q_A(\mathbf{x})$ (and, in turn, the quadratic approximation of f at $(0,0)$) corresponds to plot (1).

(B) Since g has a critical point at $(0,0)$, the quadratic approximation of g at the origin takes the form

$$\begin{aligned} g(h,k) &\approx g(0,0) + \frac{1}{2}\begin{bmatrix} h & k \end{bmatrix} ((\mathrm{Hg})(0,0)) \begin{bmatrix} h \\ k \end{bmatrix} \\ &= g(0,0) + \frac{1}{2}\begin{bmatrix} h & k \end{bmatrix} \begin{bmatrix} 4 & 10 \\ 10 & 4 \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix} \\ &= g(0,0) + 2h^2 + 10hk + 2k^2 \\ &= g(0,0) + q_B(h,k), \end{aligned}$$

where B is the matrix

$$B = \begin{bmatrix} 2 & 5 \\ 5 & 2 \end{bmatrix}.$$

Level sets of g's quadratic approximation resemble (up to a re-labeling of levels) level sets of q_B. The matrix B has trace 4 and determinant $4 - 25 = -21$, so it has characteristic polynomial

$$\lambda^2 - 4\lambda - 21 = (\lambda - 7)(\lambda + 3),$$

which has as its roots $\lambda_1 = 7$ and $\lambda_2 = -3$ (which can also be found by the quadratic formula, if you didn't notice how it factors).

To sketch the contour plot, we need to work out perpendicular eigenvectors $\mathbf{w}_1$ and $\mathbf{w}_2$ for λ_1 and λ_2 respectively so as to write the quadratic form q_B in a more convenient reference frame. The lines for the eigenvalues are the null spaces of $B - 7\mathrm{I}_2$ and $B + 3\mathrm{I}_2$. We compute these matrices to be

$$B - 7\mathrm{I}_2 = \begin{bmatrix} -5 & 5 \\ 5 & -5 \end{bmatrix}, \qquad B + 3\mathrm{I}_2 = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix}.$$

The first has null space corresponding to the pair of equations $-5x + 5y = 0$ and $5x - 5y = 0$, which are scalar multiples of each other (as they must be for a line of an eigenvalue): this is the line $y = x$, so it is the span of the vector $\mathbf{w}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ (or any nonzero scalar multiple of this). The second of these has null space corresponding to the pair of equations $5x + 5y = 0$ and $5x + 5y = 0$, which are likewise scalar multiples of each other (as they must be for a line of an eigenvalue): this is the line $y = -x$, so it is the span of the vector $\mathbf{w}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ (or any nonzero scalar multiple of this). The quadratic form q_B is given by the formula

$$q_B(c_1\mathbf{w}_1 + c_2\mathbf{w}_2) = \lambda_1(\mathbf{w}_1 \cdot \mathbf{w}_1)c_1^2 + \lambda_2(\mathbf{w}_2 \cdot \mathbf{w}_2)c_2^2 = 7(2)c_1^2 - 3(2)c_2^2.$$

This has level curves that are hyperbolas centered at $(0, 0)$ with asymptotes closer to the line spanned by $\mathbf{w}_2$ since $|\lambda_2| < |\lambda_1|$.

Thus, $q_B(\mathbf{x})$ (and, in turn, the quadratic approximation of g at $(0, 0)$) corresponds to plot (4).

4. (2 points) A researcher collected 100 data points

$$(x_1, y_1),\ (x_2, y_2),\ \ldots,\ (x_{100}, y_{100})$$

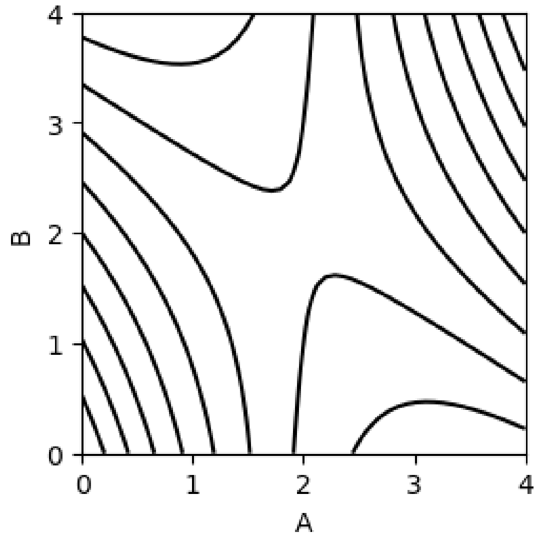
and attempted to find the best-fit line $y = A + Bx$; i.e., she wanted to find A and B that minimizes the total squared-error function

$$f(A, B) = (y_1 - A - Bx_1)^2 + (y_2 - A - Bx_2)^2 + \cdots + (y_{100} - A - Bx_{100})^2.$$

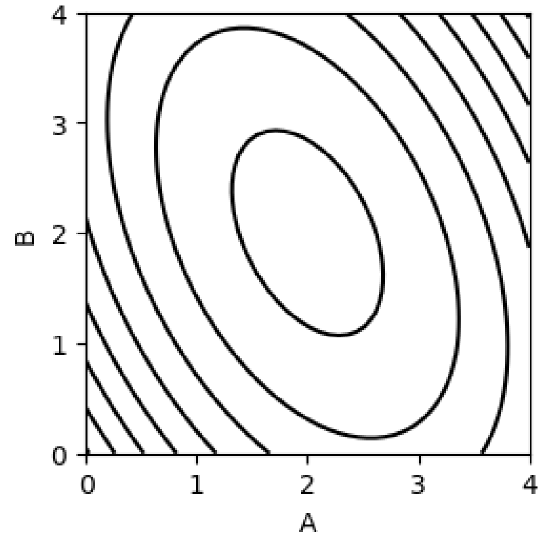
Note that $f(A, B)$ is a quadratic function of two variables A, B.

Which of the following are possible contour plots of f (with horizontal A-axis and vertical B-axis)? Select all that apply.

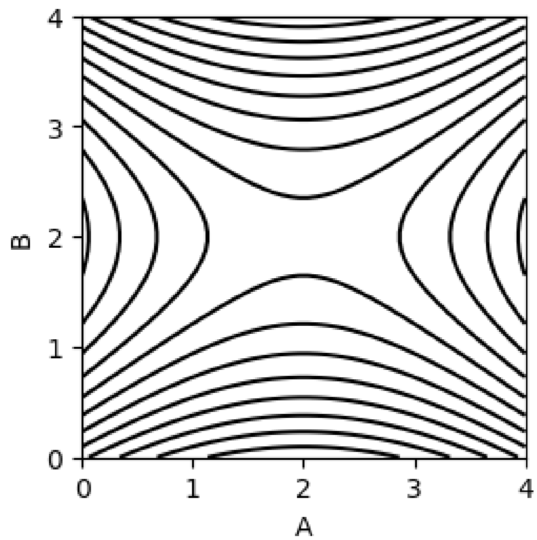
(i)



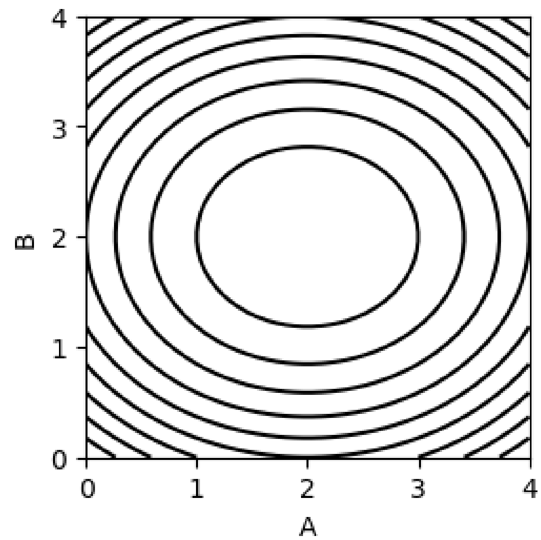
(ii)



(iii)



(iv)



Among the choices, the correct answers are (ii) and (iv).

Recall that if $\mathbf{X}$ and $\mathbf{Y}$ are the 100-vectors consisting of data-point values $x_1, \ldots, x_{100}$ and $y_1, \ldots, y_{100}$ respectively, then the function

$$f(A, B) = \|\mathbf{Y} - (A\mathbf{1} + B\mathbf{X})\|^2,$$

where $\mathbf{1}$ is the 100-vector consisting of all 1's. We know that this function is minimized when variables A, B take the values $(A, B) = (A_0, B_0)$ for which

$$(*) \qquad A_0\mathbf{1} + B_0\mathbf{X} = \mathbf{Proj}_{\mathrm{span}(\mathbf{1},\mathbf{X})}(\mathbf{Y});$$

but in terms of calculus, we may say that this (A_0, B_0) is a *critical point* for $f(A, B)$.

Near such a critical point, the function f's quadratic approximation (which is equal to f itself, since f is quadratic) has the same level sets (up to a relabeling of level values) as those of a quadratic form, centered at the critical point (A_0, B_0) of f. Since the formula for f is clearly non-negative for any (A, B) (it is a squared length, or a sum of squared-differences), seeing level sets of *indefinite* quadratic forms (i.e., hyperbolas), such as in choices (i) or (iii), are not possible. On the other hand, level sets of *positive-definite* quadratic forms (i.e., ellipses), such as in choices (ii) and (iv), are certainly possible.

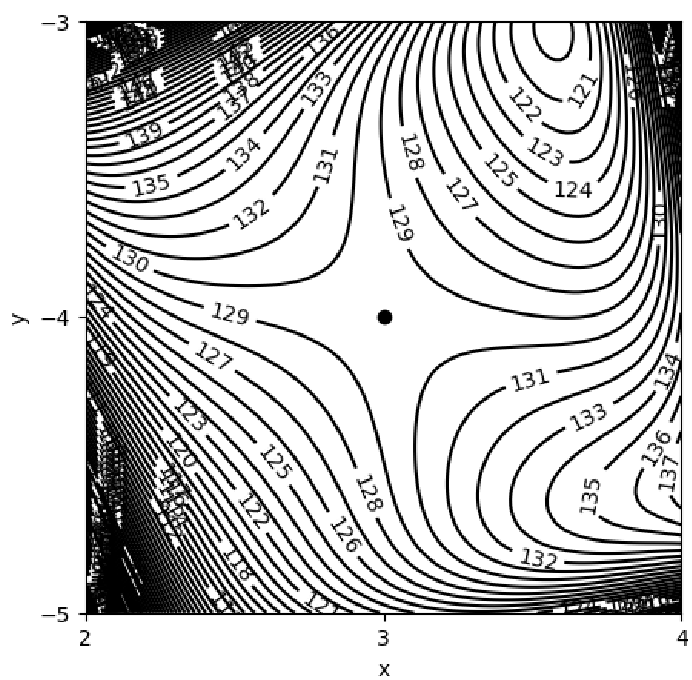
Remark 1: The particular tilt-angle and length-ratio of the axes of the ellipses arising in the level curves of $f(A, B)$ will depend on the particular 100 data-point values. For example, if the data is such that $\mathbf{X} \cdot \mathbf{1} = 0$ (x-values are "centered," or have mean zero), then

$$\begin{aligned} f(A,B) &= \|\mathbf{Y} - (A\mathbf{1} + B\mathbf{X})\|^2 \\ &= (\mathbf{Y}\cdot\mathbf{Y}) - 2A(\mathbf{Y}\cdot\mathbf{1}) - 2B(\mathbf{Y}\cdot\mathbf{X}) + A^2(\mathbf{1}\cdot\mathbf{1}) + B^2(\mathbf{X}\cdot\mathbf{X}) \end{aligned}$$

which at the critical point will have the same level sets as those of a quadratic form with no cross-term (coefficient of AB); these ellipse axes will be aligned to the A-, B-coordinate axes, with length ratio dependent on the ratio $\frac{\|\mathbf{X}\|}{\|\mathbf{1}\|}$. Other tilt-angles can arise for different values of the 100 data points.

Remark 2: Ellipses (or circles) arising as level sets near the critical point of $f(A, B)$ are practically the *only* picture one can expect to see — unless the original 100 data points already lie on a vertical line! Remember that by properties of projections, $\mathbf{Proj}_{\mathrm{span}(\mathbf{1},\mathbf{X})}(\mathbf{Y})$ is the *unique* vector in span$(\mathbf{1}, \mathbf{X})$ with this closest distance to $\mathbf{Y}$; so if $\mathbf{1}, \mathbf{X}$ are not scalings of each other (always true unless all 100 data points have the same x-value), then there will in turn be a unique pair (A_0, B_0) that expresses the projection as a combination of $\mathbf{1}$ and $\mathbf{X}$ as in $(*)$; this means that (A_0, B_0) will be the unique minimum point of f, and the non-negative quadratic form seen in f will in fact be positive-definite (and not simply positive-semidefinite). Positive-definite quadratic forms in two variables always yield ellipse-shaped (or circle-shaped) level sets.

5. (4 points) Below is a contour plot of $f(x, y)$ around a critical point $(3, -4)$:



At this critical point, the Hessian matrix $(\mathrm{Hf})(3, -4)$ has one positive eigenvalue, λ, and one negative eigenvalue, μ.

(a) Which of the following is approximately equal to an eigenvector of $(\mathrm{Hf})(3, -4)$ with positive eigenvalue (i.e., with eigenvalue λ)?

(b) Which of the following is approximately equal to an eigenvector of $(\mathrm{Hf})(3, -4)$ with negative eigenvalue (i.e., with eigenvalue μ)?

(i) $(1, 1)$ (ii) $(1, 0)$ (iii) $(-1, 1)$ (iv) $(0, 1)$

From the picture, $f(3,-4)$ lies between 129 and 130, and level sets near $(3,-4)$ are approximate hyperbolas. Furthermore, near this critical point the quadratic approximation of f (for small h, k) is given by

$$\begin{aligned} f(3+h,-4+k) &\approx f(3,-4) + \frac{1}{2}\begin{bmatrix} h & k \end{bmatrix}\Big((\mathrm{Hf})(3,-4)\Big)\begin{bmatrix} h \\ k \end{bmatrix} \\ &= f(3,-4) + \frac{1}{2}q_H(h,k) \end{aligned}$$

where $q_H = q_{(\mathrm{Hf})(3,-4)}$ is the quadratic form associated to $(\mathrm{Hf})(3,-4)$. This tells us that for sufficiently small h, k,

- $f(3+h,-4+k) < f(3,-4)$ (roughly) corresponds to $q_H(h,k) < 0$; and
- $f(3+h,-4+k) > f(3,-4)$ (roughly) corresponds to $q_H(h,k) > 0$.

We also know that eigenvectors of $(\mathrm{Hf})(3,-4)$ span the "principal axes" of the level sets of its associated quadratic form; that is to say, if in particular the level sets of q_H are hyperbolas (as they are here, since the level sets of f are approximate hyperbolas as we noted), then eigenvectors lie along the (perpendicular) axes of bilateral symmetry of these hyperbola families. In the diagram, these axes appear to be approximately of slope ± 1, so $\boxed{\text{options (i) and (iii)}}$ are candidate answers for parts (a) and (b).

Finally, whenever $\mathbf{u}$ is an eigenvector of some symmetric matrix A with associated eigenvalue α, then

$$q_A(\mathbf{u}) = \mathbf{u}\cdot A\mathbf{u} = \mathbf{u}\cdot(\alpha\mathbf{u}) = \alpha(\mathbf{u}\cdot\mathbf{u}) = \alpha\|\mathbf{u}\|^2,$$

and in particular we can say that α and $q_A(\mathbf{u})$ always have the same sign whenever either is nonzero.

(a) By the above, an eigenvector $\mathbf{v}$ of $(\mathrm{Hf})(3,-4)$ with positive eigenvalue will have $q_H(\mathbf{v}) > 0$, so for $\mathbf{v} = (h,k)$ of sufficiently small length, we expect to see $f(3+h,-4+k) > f(3,-4)$. Among the candidate directions $(1,1)$ and $(-1,1)$ (i.e., small scalar multiples of these), we see from the picture this is consistent with $(-1,1)$, or $\boxed{\text{choice (iii)}}$. (That is, points slightly above-left and below-right of $(3,-4)$ have f-values greater than $f(3,-4)$.)

(b) Similarly, an eigenvector $\mathbf{w}$ of $(\mathrm{Hf})(3,-4)$ with negative eigenvalue will have $q_H(\mathbf{w}) < 0$, so for $\mathbf{w} = (h,k)$ of sufficiently small length, we expect to see $f(3+h,-4+k) < f(3,-4)$. Among the candidate directions $(1,1)$ and $(-1,1)$ (i.e., small scalar multiples of these), we see from the picture this is consistent with $(1,1)$, or $\boxed{\text{choice (i)}}$. (That is, points slightly above-right and below-left of $(3,-4)$ have f-values less than $f(3,-4)$.)

6. (2 points) **True or False:** Suppose $f(x,y)$ has the property that $\frac{\partial^2 f}{\partial x^2}(x,y) + \frac{\partial^2 f}{\partial y^2}(x,y) = 1$ for all (x,y). Then f has no local maxima.

The statement is $\boxed{\text{true}}$. In order for f to have a local maximum, the Hessian matrix of f must be at least negative semi-definite, i.e. the associated quadratic form cannot attain any positive values. However, the second derivatives $\frac{\partial^2 f}{\partial x^2}$ and $\frac{\partial^2 f}{\partial y^2}$ are the diagonal entries of the Hessian matrix, and the fact that they always add up to 1 means that at least one of them must be positive for any (x,y). This means the Hessian matrix can never be negative semi-definite, since $\begin{bmatrix} 1 \\ 0 \end{bmatrix}^T Hf \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{\partial^2 f}{\partial x^2}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}^T Hf \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{\partial^2 f}{\partial y^2}$, and at least one of these values must be positive at any point.

7. (2 points) **True or False:** Suppose $f(x,y)$ has the property that $Hf(x,y)$ is positive definite for any point (x,y), and let $g(x,y) = e^{f(x,y)}$. Then $Hg(x,y)$ is also positive definite for any point (x,y).

The statement is $\boxed{\text{true}}$. We aim to compute the Hessian matrix of g. From the chain rule, we have

$$\frac{\partial g}{\partial x}(x,y) = e^{f(x,y)}\frac{\partial f}{\partial x}(x,y) \quad \text{and} \quad \frac{\partial g}{\partial y}(x,y) = e^{f(x,y)}\frac{\partial f}{\partial y}(x,y),$$

from which we can compute using the product rule that

$$\frac{\partial^2 g}{\partial x^2}(x,y) = \frac{\partial}{\partial x}(e^{f(x,y)})\frac{\partial f}{\partial x}(x,y) + e^{f(x,y)}\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}(x,y)\right) = e^{f(x,y)}\left(\frac{\partial f}{\partial x}(x,y)\right)^2 + e^{f(x,y)}\frac{\partial^2 f}{\partial x^2}(x,y).$$

Similarly, we obtain

$$\frac{\partial^2 g}{\partial y^2}(x,y) = e^{f(x,y)}\left(\frac{\partial f}{\partial y}(x,y)\right)^2 + e^{f(x,y)}\frac{\partial^2 f}{\partial y^2}(x,y).$$

Finally, for the mixed second derivative, we have

$$\frac{\partial^2 g}{\partial x\partial y}(x,y) = \frac{\partial}{\partial x}(e^{f(x,y)})\frac{\partial f}{\partial y}(x,y)+e^{f(x,y)}\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}(x,y)\right) = e^{f(x,y)}\frac{\partial f}{\partial x}(x,y)\frac{\partial f}{\partial y}(x,y)+e^{f(x,y)}\frac{\partial^2 f}{\partial x\partial y}(x,y).$$

Noting that we can pull out a factor of $e^{f(x,y)}$ from all of these terms, we have

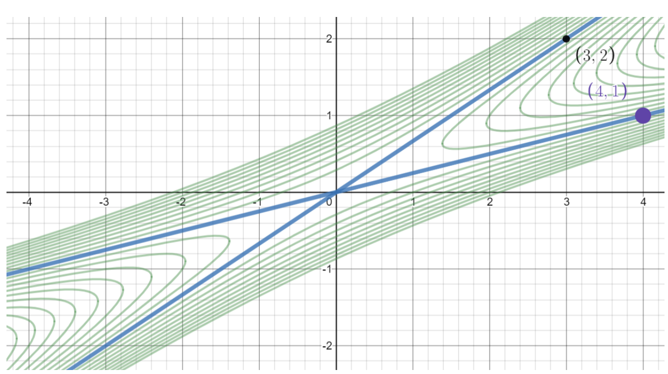
$$Hg(x,y) = e^{f(x,y)}\begin{bmatrix} \frac{\partial^2 f}{\partial x^2} + \left(\frac{\partial f}{\partial x}\right)^2 & \frac{\partial^2 f}{\partial x\partial y} + \frac{\partial f}{\partial x}\frac{\partial f}{\partial y} \\ \frac{\partial^2 f}{\partial x\partial y} + \frac{\partial f}{\partial x}\frac{\partial f}{\partial y} & \frac{\partial^2 f}{\partial y^2} + \left(\frac{\partial f}{\partial y}\right)^2 \end{bmatrix}.$$

Letting $A = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x\partial y} \\ \frac{\partial^2 f}{\partial x\partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$ and $B = \begin{bmatrix} \left(\frac{\partial f}{\partial x}\right)^2 & \frac{\partial f}{\partial x}\frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial x}\frac{\partial f}{\partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$, so that $Hg = e^f(A+B)$, we see that A is just the Hessian matrix of f, which is positive definite by assumption. On the other hand, B is positive semi-definite: indeed, we can write B as $B = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}\begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} = \nabla f(\nabla f)^T$, so $\mathbf{x}^T B\mathbf{x} = \mathbf{x}^T\nabla f(\nabla f)^T\mathbf{x} = (\mathbf{x}\cdot\nabla f)(\nabla f\cdot\mathbf{x}) = (\nabla f\cdot\mathbf{x})^2 \geq 0$. Hence, for $\mathbf{x} \neq \mathbf{0}$ we have

$$\mathbf{x}^T Hg\mathbf{x} = e^f\mathbf{x}^T A\mathbf{x} + e^f\mathbf{x}^T B\mathbf{x} \geq e^f\mathbf{x}^T A\mathbf{x} > 0,$$

i.e. Hg is positive definite.

8. (2 points) **True or False:** Consider the following partial contour plot of some quadratic form:



The blue contour corresponds to the level set at the level 0.

Then: $\begin{bmatrix}4\\1\end{bmatrix}$ and $\begin{bmatrix}3\\2\end{bmatrix}$ are eigenvectors of the matrix associated to this quadratic form.

The statement is $\boxed{\text{false}}$. The vectors $\begin{bmatrix}4\\1\end{bmatrix}$ and $\begin{bmatrix}3\\2\end{bmatrix}$ span the asymptote lines for the hyperbolas given by the level sets of this quadratic form; the asymptote lines are never the same as the eigenlines (i.e. the lines containing the eigenvectors) for the associated matrix (in fact the eigenlines bisect the angle between the asymptote lines).

Another way to see that this statement is false is that the eigenvalues associated to the matrix must be different, since the fact that some level sets are hyperbolas means the quadratic form is indefinite and thus has one positive and one negative eigenvalue, and hence the eigenvectors associated to the two different eigenvalues must be orthogonal (since the corresponding matrix is symmetric), whereas the two vectors given are not orthogonal.

9. (3 points) For which value of a is the line $y = mx$ an eigenline for the Hessian matrix of the function $f(x, y) = y^2 - axy$ at $(0, 0)$?

   (a) 0
   (b) $m/2$
   (c) m
   (d) $2m$
   (e) m^2
   (f) $2m^2 - 1$
   (g) $\frac{1}{m}$
   (h) $\frac{2m}{1-m^2}$
   (i) $\frac{m}{1+m^2}$

The Hessian of f is given by $Hf(x, y) = \begin{bmatrix}0 & -a\\-a & 2\end{bmatrix}$ (independent of x and y). Noting that $\begin{bmatrix}1\\m\end{bmatrix}$ is a vector which spans the line $y = mx$, we see that for the line to be an eigenline of Hf, we need $\begin{bmatrix}0 & -a\\-a & 2\end{bmatrix}\begin{bmatrix}1\\m\end{bmatrix} = \lambda\begin{bmatrix}1\\m\end{bmatrix}$ for some number λ. Multiplying the product on the left, we thus get

$$\begin{bmatrix}-am\\-a+2m\end{bmatrix} = \begin{bmatrix}\lambda\\\lambda m\end{bmatrix}.$$

Thus, we have $\lambda = -am$, in which case the second equation becomes $-a + 2m = \lambda m = -am^2$. We thus have $2m = (1 - m^2)a$, so solving for a yields $\boxed{a = \dfrac{2m}{1-m^2}}$.

Remark: Note that if $m = \tan\theta$, then the double-angle formula would give $a = \tan(2\theta)$. Note as well that $f(x, y) = y(y - ax)$, so the level set of f at 0 consists of the two lines $y = 0$ and $y = ax$; these two lines end up giving the asymptotes for the hyperbolas giving the level sets of f. Geometrically we expect the eigenline of the Hessian to bisect the asymptotic lines, so if an eigenline is given by $y = mx$, then this eigenline makes an angle of θ with the line $y = 0$ if we let $m = \tan\theta$; hence it would make geometric sense that the angle between the asymptotes would be 2θ, i.e. that $a = \tan(2\theta)$.

10. (10 points) Find and classify all critical points of $f(x, y) = (x + y - 2)^2 + (x - y)^3 - (x - y)^2$.

We compute

$$\frac{\partial f}{\partial x} = 2(x+y-2) + 3(x-y)^2 - 2(x-y) \quad \text{and} \quad \frac{\partial f}{\partial y} = 2(x+y-2) - 3(x-y)^2 + 2(x-y).$$

Hence, to compute the critical points of f, we need to solve the equations $2(x+y-2)+3(x-y)^2-2(x-y) = 0$ and $2(x+y-2)-3(x-y)^2+2(x-y)$. Adding the two equations gives $4(x+y-2) = 0$, and hence $x+y-2 = 0$. Substituting $y = 2-x$, the first equation then becomes $3(x-(2-x))^2-2(x-(2-x)) = 0$, i.e. $3(2x-2)^2 - 2(2x-2) = 0$. The quadratic factors as $(2x-2)(3(2x-2)-2) = (2x-2)(6x-8)$. Thus, either $2x-2 = 0 \implies x = 1$ (which then gives $y = 1$), or $6x-8 = 0 \implies x = 4/3$ (which then gives $y = 2/3$). Hence, the critical points are $(1, 1)$ and $(4/3, 2/3)$.

Furthermore, we can compute the second derivatives as follows:

$$\frac{\partial^2 f}{\partial x^2} = 2+6(x-y)-2 = 6(x-y), \quad \frac{\partial^2 f}{\partial y^2} = 2+6(x-y)-2 = 6(x-y), \quad \frac{\partial^2 f}{\partial x \partial y} = 2-6(x-y)+2 = 4-6(x-y).$$

Hence we have

$$Hf(x, y) = \begin{bmatrix} 6(x-y) & 4-6(x-y) \\ 4-6(x-y) & 6(x-y) \end{bmatrix}.$$

At the critical point $(1, 1)$, we have

$$Hf(1, 1) = \begin{bmatrix} 0 & 4 \\ 4 & 0 \end{bmatrix}.$$

The characteristic polynomial of the above matrix is $\lambda^2 - 16$, which has roots ± 4, so the Hessian matrix is indefinite. Hence, the critical point $(1, 1)$ is a saddle point.

At the critical point $(4/3, 2/3)$, we have

$$Hf(4/3, 2/3) = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}.$$

This matrix is evidently positive definite (it is 4 times the identity matrix), so the critical point $(4/3, 2/3)$ is a local minimum.

Remark 1: This function appeared in the check-in question on PCRQ 10. In that question, you were expected to identify $(1, 1)$ as a saddle point without using the machinery from Chapters 25 and 26 just by analyzing the behavior of the function along the lines parametrized by $(1+t, 1+t)$ and $(1+t, 1-t)$. Note that the vectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ are tangent to these lines; these vectors also happen to be the eigenvectors for the Hessian matrix at that point, as you can verify by inspection.

Remark 2: The local minimum at $(4/3, 2/3)$ turns out not to be a global minimum; in fact there turns out to be no global extrema of either kind. Indeed, along the line $(1+t, 1-t)$ we have

$$f(1+t, 1-t) = 8t^3 - 4t^2,$$

so taking $t \to +\infty$ will make f take on arbitrary large positive values, while taking $t \to -\infty$ will make f take on arbitrarily large negative values.