Solutions to Math 51 Final Exam (Practice #2)

1. (10 points) Find all points $P = (a, b)$ on the curve

$$x^2 + \frac{2y^3}{3} = 3$$

that are closest to the origin (equivalently: points on this curve that minimize $x^2 + y^2$), as well as their common minimal *distance* to the origin.

Define $f(x, y) = x^2 + y^2$ and $g(x, y) = x^2 + (2/3)y^3$, so we seek to minimize f subject to the condition $g = 3$. At a point $P = (a, b)$ with minimal distance we have either $(\nabla g)(P) = \mathbf{0}$ or $(\nabla f)(P) = \lambda(\nabla g)(P)$ for some scalar λ.

Computing the gradients symbolically, we have

$$(\nabla g)(x, y) = \begin{bmatrix} 2x \\ 2y^2 \end{bmatrix}, \ (\nabla f)(x, y) = \begin{bmatrix} 2x \\ 2y \end{bmatrix}.$$

The only point at which the gradient of g vanishes is the origin, but that does not lie on the constraint curve $g = 3$. Hence, at $P = (a, b)$ we must have

$$\begin{bmatrix} 2a \\ 2b \end{bmatrix} = \lambda \begin{bmatrix} 2a \\ 2b^2 \end{bmatrix}.$$

Comparing vector entries, this says

$$2a = 2a\lambda, \ \ 2b = 2b^2\lambda.$$

To carry out the usual method of comparing various fractional expressions for λ, we need to first identify where the various entries in $(\nabla g)(a, b)$ vanish: $a = 0$ or $b = 0$. If $b = 0$ then the constraint $3 = g(a, b) = g(a, 0) = a^2$ forces $a = \pm\sqrt{3}$, so we will come back to the points $(\pm\sqrt{3}, 0)$ later. Thus, we may assume $b \neq 0$. If $a = 0$ then the constraint $3 = g(0, b) = (2/3)b^3$ forces $b = (9/2)^{1/3}$, so we will come back to the point $(0, (9/2)^{1/3})$ later.

Now assuming $a, b \neq 0$, we can divide by each, so the first equation "$2a = 2a\lambda$" gives $\lambda = 1$ and then the second equation "$2b = 2b^2\lambda$" becomes $2b = 2b^2$, which is to say $b = 1$ (since $b \neq 0$). But then going back to the constraint gives $3 = g(a, b) = g(a, 1) = a^2 + (2/3)$, so $a^2 = 7/3$ or equivalently $a = \pm\sqrt{7/3}$.

Putting it all together, we have several candidate points on the curve $g = 3$:

$$(\pm\sqrt{7/3}, 1), \ \ (\pm\sqrt{3}, 0), \ \ (0, (9/2)^{1/3}).$$

We have shown that the minimizers for $f(x, y) = x^2 + y^2$ are among such points, so we compute f on each and compare the outcomes:

$$f(\pm\sqrt{7/3}, 1) = (7/3) + 1 = 10/3, \ \ f(\pm\sqrt{3}, 0) = 3, \ \ f(0, (9/2)^{1/3}) = (9/2)^{2/3}.$$

Clearly $10/3 > 3$, so we just need to figure out which among 3 and $(9/2)^{2/3}$ is smaller to determine which among $(\pm\sqrt{3}, 0)$ and $(0, (9/2)^{1/3})$ is nearest the origin. It is the same to compare these numbers after cubing them, which is to say: which among 27 and $(9/2)^2 = 81/4$ is smaller. But $81/4 = 20.25$, so it is smaller. Hence, $(0, (9/2)^{1/3})$ is the unique point closest to the origin, and its distance to the origin is $(9/2)^{1/3}$.

2. (10 points) Let $f(x, y) = (xy^5 - 2y, x^2y^3 + 3x + 2y)$, so $f(2, -1) = (0, 0)$.

(a) (5 points) Compute the derivative matrix $(D(f \circ f))(2, -1)$. (Hint: do not try to compute $f \circ f$ explicitly; that is a tremendous mess. Instead, use the Chain Rule.)

By the Chain Rule, we have a matrix identity

$$(D(f \circ f))(2, -1) = (Df)(f(2, -1))\,(Df)(2, -1) = (Df)(0, 0)\,(Df)(2, -1).$$

To compute the two matrices on the right side that are being multiplied, we compute partial derivatives of the component functions of f:

$$(Df)(x, y) = \begin{bmatrix} y^5 & 5xy^4 - 2 \\ 2xy^3 + 3 & 3x^2y^2 + 2 \end{bmatrix}.$$

Hence,

$$(Df)(0, 0) = \begin{bmatrix} 0 & -2 \\ 3 & 2 \end{bmatrix}, \quad (Df)(2, -1) = \begin{bmatrix} -1 & 8 \\ -1 & 14 \end{bmatrix},$$

so

$$(D(f \circ f))(2, -1) = \begin{bmatrix} 0 & -2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} -1 & 8 \\ -1 & 14 \end{bmatrix} = \begin{bmatrix} 2 & -28 \\ -5 & 52 \end{bmatrix}.$$

(b) (5 points) Use your answer to part (a) to find the linear approximation to $(f \circ f)(2 + h, -1 + k)$ for h, k near 0. Use this to estimate $(f \circ f)(2.1, -0.95)$.

Letting $g = f \circ f : \mathbf{R}^2 \to \mathbf{R}^2$, the linear approximation to $g(2 + h, -1 + k)$ is $g(2, -1) + ((Dg)(2, -1)) \begin{bmatrix} h \\ k \end{bmatrix}$. Since $g(2, -1) = f(f(2, -1)) = f(0, 0) = (0, 0)$, by part (a) we have

$$g(2 + h, -1 + k) \approx ((Dg)(2, -1)) \begin{bmatrix} h \\ k \end{bmatrix} = \begin{bmatrix} 2 & -28 \\ -5 & 52 \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix} = \begin{bmatrix} 2h - 28k \\ -5h + 52k \end{bmatrix}.$$

To estimate $(f \circ f)(2.1, -0.95)$ we use $h = 0.1$ and $k = 0.05$ to get the estimate

$$\begin{bmatrix} 0.2 - 28(0.05) \\ -0.5 + 52(0.05) \end{bmatrix} = \begin{bmatrix} 0.2 - 2.8(0.5) \\ -0.5 + (5.2)(0.5) \end{bmatrix} = \begin{bmatrix} 0.2 - 1.4 \\ -0.5 + 2.6 \end{bmatrix} = \begin{bmatrix} -1.2 \\ 2.1 \end{bmatrix}.$$

3. (10 points) The matrix $A = \begin{bmatrix} -3 & 3 & 9 \\ -2 & 3 & 8 \\ 1 & 2 & 5 \end{bmatrix}$ is equal to LU with $L = \begin{bmatrix} 3 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 3 & 2 \end{bmatrix}$ and $U = \begin{bmatrix} -1 & 1 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$.

(a) (4 points) For $\mathbf{b} = \begin{bmatrix} 3 \\ 7 \\ -2 \end{bmatrix}$, use the given LU decomposition to solve $A\mathbf{x} = \mathbf{b}$ via repeated back-substitutions, and then check directly that your solution really is a solution. (All entries in the solution vector $\mathbf{x}$ are integers.)

We first solve $L\mathbf{y} = \mathbf{b}$, and then $U\mathbf{x} = \mathbf{y}$. The lower triangular system of equations is

$$3y_1 = 3, \;\; 2y_1 + y_2 = 7, \;\; -y_1 + 3y_2 + 2y_3 = -2.$$

This gives $y_1 = 1$, then $2 + y_2 = 7$, so $y_2 = 5$, and finally $-1 + 3(5) + 2y_3 = -2$, so $y_3 = -8$. Next, the upper triangular system is

$$-x_1 + x_2 + 3x_3 = 1, \;\; x_2 + 2x_3 = 5, \;\; x_3 = -8,$$

so going backwards gives $x_3 = -8$, so $x_2 = 21$, so $-x_1 + 21 + 3(-8) = 1$, yielding $x_1 = -4$. Hence, $\mathbf{x} = \begin{bmatrix} -4 \\ 21 \\ -8 \end{bmatrix}$. This really works, since $-3(-4) + 3(21) + 9(-8) = 12 + 63 - 72 = 3$, $-2(-4) + 3(21) + 8(-8) = 8 + 63 - 64 = 7$, and $1(-4) + 2(21) + 5(-8) = -4 + 42 - 40 = -2$.

(b) (6 points) Use the given LU decomposition to compute A^{-1} (its entries are integers or fractions with denominator at most 6), and check that what you obtain really is an inverse to A by multiplying it against A in some order (you do not need to compute the matrix product in both orders; It is recommended to check your calculations of U^{-1} and L^{-1} really work before computing A^{-1}).

Since $A = LU$ with L and U each invertible (due to having no 0 in their diagonals), we have $A^{-1} = U^{-1}L^{-1}$. To calculate U^{-1} and L^{-1}, we set them up as

$$U^{-1} = \begin{bmatrix} -1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}, \quad L^{-1} = \begin{bmatrix} 1/3 & 0 & 0 \\ a' & 1 & 0 \\ b' & c' & 1/2 \end{bmatrix}$$

and need to solve for a, b, c and a', b', c'.
For U^{-1}, we use the requirement

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = UU^{-1} = \begin{bmatrix} -1 & 1 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -a+1 & -b+c+3 \\ 0 & 1 & c+2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Looking just above the diagonal, we have $a = 1$ and $c = -2$. Then looking in the upper-right, we get $0 = -b + c + 3 = -b - 2 + 3 = -b + 1$, so $b = 1$.
Next, for L^{-1} we use the requirement

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = LL^{-1} = \begin{bmatrix} 3 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 3 & 2 \end{bmatrix} \begin{bmatrix} 1/3 & 0 & 0 \\ a' & 1 & 0 \\ b' & c' & 1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2/3 + a' & 1 & 0 \\ -1/3 + 3a' + 2b' & 3 + 2c' & 1 \end{bmatrix}.$$

Looking just below the diagonal, we have $a' = -2/3$ and $c' = -3/2$. Going into the lower-left corner, $0 = -1/3 + 3a' + 2b' = -1/3 - 2 + 2b'$, so $b' = 7/6$.
Having computed U^{-1} and L^{-1}, we multiply them (in the correct order!) to obtain A^{-1}:

$$A^{-1} = U^{-1}L^{-1} = \begin{bmatrix} -1 & 1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1/3 & 0 & 0 \\ -2/3 & 1 & 0 \\ 7/6 & -3/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1/6 & -1/2 & 1/2 \\ -3 & 4 & -1 \\ 7/6 & -3/2 & 1/2 \end{bmatrix}.$$

To check this works, we compute the matrix product

$$\begin{aligned} \begin{bmatrix} -3 & 3 & 9 \\ -2 & 3 & 8 \\ 1 & 2 & 5 \end{bmatrix} \begin{bmatrix} 1/6 & -1/2 & 1/2 \\ -3 & 4 & -1 \\ 7/6 & -3/2 & 1/2 \end{bmatrix} &= \begin{bmatrix} -1/2 - 9 + 21/2 & 3/2 + 12 - 27/2 & -3/2 - 3 + 9/2 \\ -1/3 - 9 + 28/3 & 1 + 12 - 12 & -1 - 3 + 4 \\ 1/6 - 6 + 35/6 & -1/2 + 8 - 15/2 & 1/2 - 2 + 5/2 \end{bmatrix} \\ &= \begin{bmatrix} 20/2 - 9 & -24/2 + 12 & 6/2 - 3 \\ 27/3 - 9 & 1 & 0 \\ 36/6 - 6 & -16/2 + 8 & 6/2 - 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \end{aligned}$$

as desired.

4. (10 points) Consider the matrix $A = \begin{bmatrix} 2 & -2 & -9 \\ -1 & 5 & -4 \\ 3 & -11 & 14 \end{bmatrix}$ whose columns from left to right are the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -2 \\ 5 \\ -11 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -9 \\ -4 \\ 14 \end{bmatrix}.$$

(a) (5 points) Carry out the Gram–Schmidt process for $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ to construct an orthogonal basis $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ of $\mathbf{R}^3$. Verify directly that the $\mathbf{w}_j$'s you compute are perpendicular to each other. (The entries in the $\mathbf{w}_j$'s are integers, and you should have $\mathbf{w}_2 \cdot \mathbf{w}_2 = 24 = 2^2 \cdot 6$ and $\mathbf{w}_3 \cdot \mathbf{w}_3 = 21$.)

We have $\mathbf{w}_1 = \mathbf{v}_1 = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$, and

$$\begin{aligned} \mathbf{w}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{w}_1}{\mathbf{w}_1 \cdot \mathbf{w}_1}\mathbf{w}_1 = \mathbf{v}_2 - \frac{-42}{14}\mathbf{w}_1 &= \mathbf{v}_2 + 3\mathbf{w}_1 \\ &= \begin{bmatrix} -2 \\ 5 \\ -11 \end{bmatrix} + \begin{bmatrix} 6 \\ -3 \\ 9 \end{bmatrix} \\ &= \begin{bmatrix} 4 \\ 2 \\ -2 \end{bmatrix}, \end{aligned}$$

(This satisfies $\mathbf{w}_2 \cdot \mathbf{w}_2 = 24$ as promised.) Finally,

$$\begin{aligned} \mathbf{w}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{w}_1}{\mathbf{w}_1 \cdot \mathbf{w}_1}\mathbf{w}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{w}_2}{\mathbf{w}_2 \cdot \mathbf{w}_2}\mathbf{w}_2 = \mathbf{v}_3 - \frac{28}{14}\mathbf{w}_1 - \frac{-72}{24}\mathbf{w}_2 &= \mathbf{v}_3 - 2\mathbf{w}_1 + 3\mathbf{w}_2. \\ &= \begin{bmatrix} -9 \\ -4 \\ 14 \end{bmatrix} - \begin{bmatrix} 4 \\ -2 \\ 6 \end{bmatrix} + \begin{bmatrix} 12 \\ 6 \\ -6 \end{bmatrix} \\ &= \begin{bmatrix} -1 \\ 4 \\ 2 \end{bmatrix}. \end{aligned}$$

(This satisfies $\mathbf{w}_3 \cdot \mathbf{w}_3 = 21$ as promised, and visibly the $\mathbf{w}_j$'s are all nonzero, so the $\mathbf{v}_i$'s are linearly independent.) The dot products are $\mathbf{w}_1 \cdot \mathbf{w}_2 = 8 - 2 - 6 = 0$, $\mathbf{w}_1 \cdot \mathbf{w}_3 = -2 - 4 + 6 = 0$, and $\mathbf{w}_2 \cdot \mathbf{w}_3 = -4 + 8 - 4 = 0$.

(b) (5 points) Using your work in the previous part, express each $\mathbf{v}_j$ as a linear combination of the $\mathbf{w}_i$'s, and use that to compute the QR-decomposition of A. For the orthogonal Q and upper triangular R that you have found, compute the matrix product QR to confirm it is equal to A.

From the work in (a), we have $\mathbf{v}_1 = \mathbf{w}_1$, $\mathbf{v}_2 = -3\mathbf{w}_1 + \mathbf{w}_2$, and $\mathbf{v}_3 = 2\mathbf{w}_1 - 3\mathbf{w}_2 + \mathbf{w}_3$. Define unit vectors $\mathbf{w}'_j = \mathbf{w}_j / \|\mathbf{w}_j\|$, so $\{\mathbf{w}'_1, \mathbf{w}'_2, \mathbf{w}'_3\}$ is an orthonormal basis of $\mathbf{R}^3$. We have $\|\mathbf{w}_1\| = \sqrt{14}$, $\|\mathbf{w}_2\| = \sqrt{24} = 2\sqrt{6}$, and $\|\mathbf{w}_3\| = \sqrt{21}$, so

$$\mathbf{w}_1 = \sqrt{14}\mathbf{w}'_1, \ \ \mathbf{w}_2 = 2\sqrt{6}\mathbf{w}'_2, \ \ \mathbf{w}_3 = \sqrt{21}\mathbf{w}'_3.$$

Substituting these into the expressions for the $\mathbf{v}_i$'s in terms of the $\mathbf{w}_j$'s, we have

$$\mathbf{v}_1 = \sqrt{14}\mathbf{w}'_1, \ \ \mathbf{v}_2 = -3\sqrt{14}\mathbf{w}'_1 + 2\sqrt{6}\mathbf{w}'_2, \ \ \mathbf{v}_3 = 2\sqrt{14}\mathbf{w}'_1 - 6\sqrt{6}\mathbf{w}'_2 + \sqrt{21}\mathbf{w}'_3.$$

Putting these successive coefficients as the columns of an upper triangular matrix, we obtain

$$R = \begin{bmatrix} \sqrt{14} & -3\sqrt{14} & 2\sqrt{14} \\ 0 & 2\sqrt{6} & -6\sqrt{6} \\ 0 & 0 & \sqrt{21} \end{bmatrix}.$$

Likewise, using $\mathbf{w}'_j$ as the jth column gives an orthogonal matrix

$$Q = \begin{bmatrix} 2/\sqrt{14} & 2/\sqrt{6} & -1/\sqrt{21} \\ -1/\sqrt{14} & 1/\sqrt{6} & 4/\sqrt{21} \\ 3/\sqrt{14} & -1/\sqrt{6} & 2/\sqrt{21} \end{bmatrix}.$$

Finally, we directly multiply matrices to compute that QR is equal to

$$\begin{aligned} \begin{bmatrix} 2/\sqrt{14} & 2/\sqrt{6} & -1/\sqrt{21} \\ -1/\sqrt{14} & 1/\sqrt{6} & 4/\sqrt{21} \\ 3/\sqrt{14} & -1/\sqrt{6} & 2/\sqrt{21} \end{bmatrix} \begin{bmatrix} \sqrt{14} & -3\sqrt{14} & 2\sqrt{14} \\ 0 & 2\sqrt{6} & -6\sqrt{6} \\ 0 & 0 & \sqrt{21} \end{bmatrix} &= \begin{bmatrix} 2 & -6+4 & 4-12-1 \\ -1 & 3+2 & -2-6+4 \\ 3 & -9-2 & 6+6+2 \end{bmatrix} \\ &= \begin{bmatrix} 2 & -2 & -9 \\ -1 & 5 & -4 \\ 3 & -11 & 14 \end{bmatrix}, \end{aligned}$$

which is A as desired.

5. (10 points) **True or False**

For each of the following statements, circle either TRUE (meaning, "always true") or FALSE (meaning, "not always true"), and briefly and convincingly justify your answer. 1 point for the correct choice, and the remaining points for convincing justification.

(a) (3 points) Let $\mathbf{a}, \mathbf{x} \in \mathbf{R}^2$. Suppose $f(\mathbf{a})$ is an extremum of $f(\mathbf{x})$ subject to the constraint $g(\mathbf{x}) = 0$, then $f(\mathbf{a})$ can be a local minimum, a local maximum or a saddle point of $f(\mathbf{x})$ on $\mathbf{R}^2$.

Circle one, and justify below: TRUE $\boxed{\text{FALSE}}$

Not always true: Note that at a local minimum, maximum, or saddle point, ∇f must be equal to $\mathbf{0}$. However, at a local extremum of f subject to constraint $g(\mathbf{x}) = c$, ∇f and ∇g are only required to be parallel. For example, let $f(x, y) = x + y$, and $g(x, y) = x^2 + y^2 - 2$. $f(-1, -1) = -2$ is a local min on $g(\mathbf{x}) = 0$, but $(-1, -1)$ is not a local max/min/saddle of $f(\mathbf{x})$ on $\mathbf{R}^2$ since $(\nabla f)(-1, -1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \neq \mathbf{0}$.

(b) (3 points) Let A be an $m \times n$ matrix with linearly independent columns $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbf{R}^m$. There exists a nonzero vector $\mathbf{v} \in N(A)$, the null space of A.

Circle one, and justify below: TRUE $\boxed{\text{FALSE}}$

Not always true: Suppose there were a nonzero vector $\mathbf{v} \in N(A)$, then $A\mathbf{v} = \mathbf{0}$ implies that

$$v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + \cdots + v_n\mathbf{a}_n = \mathbf{0}$$

where not all the v_i's are 0. This means the columns $\mathbf{a}_1, \cdots, \mathbf{a}_n$ are not linearly independent, contradicting the assumption that $N(A)$ contains a nonzero vector.

(c) (4 points) Let A, B be two $n \times n$ matrices, suppose A is invertible and B is not invertible, then AB may be invertible.

Circle one, and justify below: TRUE $\boxed{\text{FALSE}}$

Not always true: AB is never invertible if B is not invertible. If AB were invertible, then $(AB)^{-1}$ exists, $((AB)^{-1}A)B = (AB)^{-1}(AB) = I_n$, so $(AB)^{-1}A$ would be an inverse of B, contradicting the condition that B is not invertible.

6. (10 points) Let $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$, and P be the 3×3 matrix representing projection onto the line spanned by $\mathbf{v}$, namely for any $\mathbf{x} \in \mathbf{R}^3$, $\mathbf{Proj_v\,x} = P\mathbf{x}$.

(a) (4 points) Write down the matrix P.

Given any $\mathbf{x} \in \mathbf{R}^3$,

$$\mathbf{Proj_v\,x} = \frac{\mathbf{x}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v} = \frac{1}{\|\mathbf{v}\|^2}\mathbf{v}\mathbf{v}^\top\mathbf{x} = \frac{1}{6}\begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \end{bmatrix}\mathbf{x}$$

So

$$P = \frac{1}{6}\begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \end{bmatrix}$$

Alternatively, one can find that

$$\mathbf{Proj_v\,e}_1 = \frac{\mathbf{e}_1\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v} = \frac{1}{6}\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{Proj_v\,e}_2 = \frac{\mathbf{e}_2\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v} = \frac{1}{6}\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{Proj_v\,e}_3 = \frac{\mathbf{e}_3\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v} = \frac{2}{6}\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

(b) (3 points) Find a basis for $C(P)$, the column space of P, and $N(P)$, the null space of P.

The columns of P are all scalar multiples of one another, so $\dim(C(P)) = 1$, and $C(P)$ has basis $\left\{\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}\right\}$.

The null space of P, $N(P) = \{\mathbf{x} \in \mathbf{R}^3 : P\mathbf{x} = \mathbf{0}\}$. Solving $P\mathbf{x} = \mathbf{0}$, we find $x_1 + x_2 + 2x_3 = 0$, so

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_2 - 2x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}.$$

A basis for $N(P)$ is $\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}\right\}$.

(c) (3 points) For each of the following $\mathbf{b}_i$, determine whether $P\mathbf{x} = \mathbf{b}_i$ has a solution. If yes, describe all solutions to the system.

a) $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ b) $\mathbf{b}_2 = \begin{bmatrix} -1 \\ -1 \\ -2 \end{bmatrix}$ c) $\mathbf{b}_3 = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$

Since $C(P)$ is spanned by $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$, only $\mathbf{b}_2 \in C(P)$, so only $P\mathbf{x} = \mathbf{b}_2$ has solution. Since $P\begin{bmatrix} -6 \\ 0 \\ 0 \end{bmatrix} =$

$\begin{bmatrix} -1 \\ -1 \\ -2 \end{bmatrix}$, and $N(P)$ has basis $\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}$, the general solution to $P\mathbf{x} = \mathbf{b}_2$ is given by

$$\begin{bmatrix} -6 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}.$$

7. (10 points) Let $f : \mathbf{R}^2 \to \mathbf{R}$ be a function satisfying

$$f(1,3) = 10, \quad (\nabla f)(1,3) = \begin{bmatrix} 3 \\ -5 \end{bmatrix}, \quad (\mathrm{H}f)(1,3) = \begin{bmatrix} 6 & -1 \\ -1 & 2 \end{bmatrix}.$$

(a) (3 points) Compute the quadratic approximation for f at $(1,3)$ (i.e., the quadratic approximation to $f(1+h, 3+k)$ for $h, k \approx 0$) and use it to estimate $f(1.2, 2.9)$.

The quadratic approximation is

$$\begin{aligned} f(1+h, 3+k) &\approx f(1,3) + (\nabla f)(1,3) \begin{bmatrix} h \\ k \end{bmatrix} + \frac{1}{2} \begin{bmatrix} h & k \end{bmatrix} (\mathrm{H}f)(1,3) \begin{bmatrix} h \\ k \end{bmatrix} \\ &= 10 + 3h - 5k + \frac{1}{2}(6h^2 - 2hk + 2k^2) \\ &= 10 + 3h - 5k + 3h^2 - hk + k^2. \end{aligned}$$

To estimate $f(1.2, 2.9)$ we use $h = 0.2$ and $k = -0.1$ to get the estimate

$$f(1.2, 2.9) \approx 10 + 0.6 + 0.5 + 3(0.04) + 0.02 + 0.01 = 10 + 1.10 + 0.12 + 0.03 = 11.10 + 0.15 = 11.25.$$

(b) (7 points) Suppose f also has a critical point at the origin, with $f_{xx}(0,0) = 7$, $f_{xy}(0,0) = 2$, and $f_{yy}(0,0) = 4$. Use eigenvalues and eigenvectors of H to determine the nature of this critical point (local maximum, local minimum, or saddle point) and to sketch the contour plot of f near $(0,0)$. (The eigenvalues of the Hessian are integers.)

The Hessian $H = (\mathrm{H}f)(0,0)$ at the origin is

$$H = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix},$$

so this 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). These are integers as promised and have the same positive sign, so q_H is positive-definite and hence $(0,0)$ is a local minimum.

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_H associated to the Hessian H in a more convenient reference frame. The lines for the eigenvalues are the null spaces of $H - 8\mathrm{I}_2$ and $H - 3\mathrm{I}_2$. We compute these matrices to be

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

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_H is given by the formula

$$q_H(u_1\mathbf{w}_1 + u_2\mathbf{w}_2) = \lambda_1(\mathbf{w}_1 \cdot \mathbf{w}_1)u_1^2 + \lambda_2(\mathbf{w}_2 \cdot \mathbf{w}_2)u_2^2 = 8(5)u_1^2 + 3(5)u_2^2.$$

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 $40u_1^2 + 15u_2^2 = c$ crosses the u_1-axis at $\pm\sqrt{c/40}$ and the u_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 sketch is shown below.

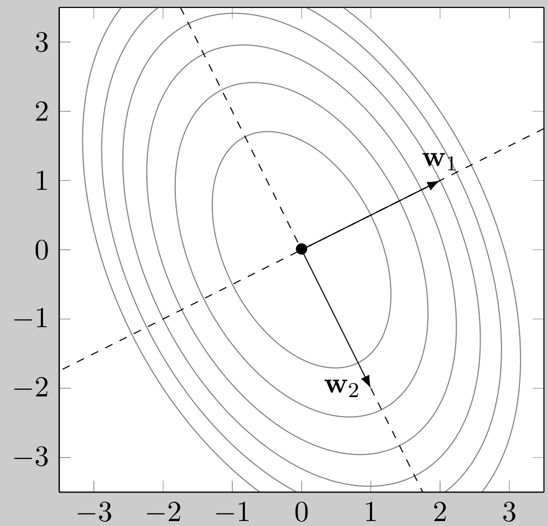

Figure 1: Approximate contour plot of f via its quadratic approximation at $(0,0)$.

8. (8 points) For each eigenvalue λ of $A = \begin{bmatrix} 5 & 0 & 0 \\ 6 & 3 & 0 \\ -4 & 2 & 2 \end{bmatrix}$, compute a basis for the nonzero linear subspace $N(A - \lambda \mathrm{I}_3)$ in $\mathbf{R}^3$, and as a check on your work verify directly that each vector in that basis is an eigenvector for A with eigenvalue λ.

As was discussed in class, the eigenvalues for a lower triangular matrix (as well as upper triangular) are its diagonal entries. So this matrix has as its eigenvalues $\lambda_1 = 5$, $\lambda_2 = 3$, and $\lambda_3 = 2$. We need to compute a basis for each of the null spaces $N(A - 5\mathrm{I}_3)$, $N(A - 3\mathrm{I}_3)$, and $N(A - 2\mathrm{I}_3)$.

For the first of these, we have

$$A - 5\mathrm{I}_3 = \begin{bmatrix} 0 & 0 & 0 \\ 6 & -2 & 0 \\ -4 & 2 & -3 \end{bmatrix},$$

so a vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is in its null space precisely when it satisfies the equations

$$0 = 0, \quad 6x_1 - 2x_2 = 0, \quad -4x_1 + 2x_2 - 3x_3 = 0.$$

The first equation says nothing, the second says $x_2 = 3x_1$, and plugging this into the third gives

$$0 = -4x_1 + 2x_2 - 3x_3 = -4x_1 + 6x_1 - 3x_3 = 2x_1 - 3x_3,$$

so $x_1 = (3/2)x_3$ and hence $x_2 = 3x_1 = (9/2)x_3$. (We could just as well solve for x_2 and x_3 in terms of x_1, for example.) This gives the null space as the collection of vectors of the form

$$\begin{bmatrix} (3/2)x_3 \\ (9/2)x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 3/2 \\ 9/2 \\ 1 \end{bmatrix},$$

so this is a line with basis given by $\mathbf{v}_1 = \begin{bmatrix} 3/2 \\ 9/2 \\ 1 \end{bmatrix}$ (or any nonzero scalar multiple of this). Direct matrix-vector multiplication gives

$$A\mathbf{v}_1 = \begin{bmatrix} 15/2 \\ 9 + 27/2 \\ -6 + 9 + 2 \end{bmatrix} = \begin{bmatrix} 15/2 \\ 45/2 \\ 5 \end{bmatrix} = 5\mathbf{v}_1,$$

as desired.

Turning to the null space of $A - 3\mathrm{I}_3$, we have

$$A - 3\mathrm{I}_3 = \begin{bmatrix} 2 & 0 & 0 \\ 6 & 0 & 0 \\ -4 & 2 & -1 \end{bmatrix}.$$

Hence, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ lies in the null space of this precisely when it satisfies the equations

$$2x_1 = 0, \;\; 6x_1 = 0, \;\; -4x_1 + 2x_2 - x_3 = 0.$$

The first two equations both say $x_1 = 0$, and plugging this into the last turns it into the condition $2x_2 - x_3 = 0$, which says $x_3 = 2x_2$. Hence,

$$\mathbf{x} = \begin{bmatrix} 0 \\ x_2 \\ 2x_2 \end{bmatrix} = x_2 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix},$$

so a basis for this null space is given by the vector $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$ (or any nonzero scalar multiple of that). Direct matrix-vector multiplication establishes the eigenvector condition:

$$A\mathbf{v}_2 = \begin{bmatrix} 0 \\ 3 \\ 2 + 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix} = 3\mathbf{v}_2$$

as desired.

Finally, to compute $N(A - 2\mathrm{I}_3)$ we have

$$A - 2\mathrm{I}_3 = \begin{bmatrix} 3 & 0 & 0 \\ 6 & 1 & 0 \\ -4 & 2 & 0 \end{bmatrix}.$$

Hence, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ lies in the null space of this precisely when it satisfies the equations

$$3x_1 = 0, \;\; 6x_1 + x_2 = 0, \;\; -4x_1 + 2x_2 = 0.$$

The first equation says $x_1 = 0$, and then each of the other two amount to the further condition $x_2 = 0$. Thus, such vectors $\mathbf{x}$ are those of the form

$$\begin{bmatrix} 0 \\ 0 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

so a basis for this null space is given by the vector $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ (or any nonzero scalar multiple of that).
This is $\mathbf{e}_3$, and the *third column* of A is $2\mathbf{e}_3$, confirming that $A\mathbf{v}_3 = 2\mathbf{v}_3$ as desired (and this can also be verified directly by matrix-vector multiplication, as for the other two cases).

9. (12 points) Consider the quadratic form $q(x, y) = 2x^2 + 2\sqrt{2}xy + 3y^2$. We want to use gradient descent to find the minimizer $\mathbf{a}$ of q where $q(\mathbf{a})$ is minimum.

(a) (4 points) Let A be the symmetric matrix associated with the quadratic form $q(x, y)$. Find A and its eigenvalues λ_1, λ_2 and their respective eigenvectors $\mathbf{v}_1, \mathbf{v}_2$.

$$A = \begin{bmatrix} 2 & \sqrt{2} \\ \sqrt{2} & 3 \end{bmatrix}$$

We find the characteristic polynomial of A to be

$$p_A(\lambda) = \det(A - \lambda I_2) = \lambda^2 - 5\lambda + 4.$$

Since $\lambda^2 - 5\lambda + 4 = (\lambda - 1)(\lambda - 4)$, the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 4$.
For $\lambda_1 = 1$, $A - I_2 = \begin{bmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 2 \end{bmatrix}$, so $N(A - I_2)$ is spanned by $\begin{bmatrix} -\sqrt{2} \\ 1 \end{bmatrix}$.
For $\lambda_2 = 4$, $A - I_2 = \begin{bmatrix} -2 & \sqrt{2} \\ \sqrt{2} & -1 \end{bmatrix}$, so $N(A - 4I_2)$ is spanned by $\begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix}$.
The corresponding eigenvectors for λ_1 and λ_2 are $\mathbf{v}_1 = \begin{bmatrix} -\sqrt{2} \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix}$ respectively.

(b) (1 point) The gradient descent algorithm to minimize $q(x, y)$ with learning rate t ($t > 0$) is given by

$$\mathbf{x}_{k+1} = \mathbf{x}_k - t(\nabla q)(\mathbf{x}_k), \qquad \text{where } \mathbf{x}_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix}.$$

Show that

$$\mathbf{x}_{k+1} = (I_2 - 2tA)\mathbf{x}_k$$

$$(\nabla q)(x, y) = \begin{bmatrix} 4x + 2\sqrt{2}y \\ 2\sqrt{2}x + 6y \end{bmatrix} = \begin{bmatrix} 4 & 2\sqrt{2} \\ 2\sqrt{2} & 6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 2A \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k - t(\nabla q)(\mathbf{x}_k) = \mathbf{x}_k - t2A\mathbf{x}_k = (I_2 - 2tA) \begin{bmatrix} x_k \\ y_k \end{bmatrix}.$$

(c) (5 points) Let $\mathbf{x}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$. Express $\mathbf{x}_0$ as a linear combination of the eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ you found in part (a), and then find $\mathbf{x}_{51} = \begin{bmatrix} x_{51} \\ y_{51} \end{bmatrix}$.

Note that A is symmetric, $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal. So we can apply Fourier's formula to find

$$\mathbf{x}_0 = \frac{\mathbf{x}_0 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{x}_0 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = \frac{-3\sqrt{2}}{3}\mathbf{v}_1 + \frac{3}{3}\mathbf{v}_2 = -\sqrt{2}\mathbf{v}_1 + \mathbf{v}_2$$

Note that

$$(I_2 - 2tA)\mathbf{v}_1 = \mathbf{v}_1 - 2tA\mathbf{v}_1 = \mathbf{v}_1 - 2t\mathbf{v}_1 = (1-2t)\mathbf{v}_1$$
$$(I_2 - 2tA)\mathbf{v}_2 = \mathbf{v}_2 - 2tA\mathbf{v}_2 = \mathbf{v}_2 - (2t)4\mathbf{v}_2 = (1-8t)\mathbf{v}_2$$

Hence,

$$\begin{aligned}
\mathbf{x}_{51} &= (I_2 - 2tA)\mathbf{x}_{50} \\
&= (I_2 - 2tA)^2\mathbf{x}_{49} \\
&\;\vdots \\
&= (I_2 - 2tA)^{51}\mathbf{x}_0 \\
&= (I_2 - 2tA)^{51}(-\sqrt{2}\mathbf{v}_1 + \mathbf{v}_2) \\
&= -\sqrt{2}(I_2 - 2tA)^{51}\mathbf{v}_1 + (I_2 - 2tA)^{51}\mathbf{v}_2 \\
&= -\sqrt{2}(1-2t)^{51}\mathbf{v}_1 + (1-8t)^{51}\mathbf{v}_2 && \text{since } A\mathbf{v} = \lambda\mathbf{v} \text{ implies } A^k\mathbf{v} = \lambda^k\mathbf{v}
\end{aligned}$$

(d) (2 points) The minimizer of $q(x,y)$ is $(0,0)$, i.e. $(0,0)$ is a global minimum for $q(x,y)$. Starting from $\mathbf{x}_0 = (3,0)$, for what learning rate $t > 0$ will gradient descent converge to the minimizer $(0,0)$?

Since

$$\mathbf{x}_k = -\sqrt{2}(1-2t)^k\mathbf{v}_1 + (1-8t)^k\mathbf{v}_2,$$

for gradient descent to converge to $(0,0)$, both $1-2t$ and $1-8t$ must be between -1 and 1. In particular, $-1 < 1-2t < 1$ implies $0 < t < 1$; $-1 < 1-8t < 1$ implies $0 < t < \frac{1}{4}$. So the learning rate t must be less than $\frac{1}{4}$.

This example illustrates the importance of keeping the learning rate t relatively small in order for gradient descent to find the intended solution.