1 Problem 1 (graded by Kangchen) - 24+1 points

(a) 4 points

The model parameter vector is $\mathbf{m} = [P, d]^T$;

(b) 4 points

$[\frac{\partial u}{\partial P}, \frac{\partial u}{\partial d}] = [\frac{d}{(x^2+y^2+d^2)^{3/2}}, \frac{P(x^2+y^2-2d^2)}{(x^2+y^2+d^2)^{5/2}}]$;

(c) 4 points

$$\begin{bmatrix} \frac{\partial^2 u}{\partial P^2} & \frac{\partial^2 u}{\partial P \partial d} \\ \frac{\partial^2 u}{\partial d \partial P} & \frac{\partial^2 u}{\partial d^2} \end{bmatrix} = \begin{bmatrix} 0 & \frac{x^2+y^2-2d^2}{(x^2+y^2+d^2)^{5/2}} \\ \frac{x^2+y^2-2d^2}{(x^2+y^2+d^2)^{5/2}} & \frac{6Pd^3-9pdx^2-9pdy^2}{(x^2+y^2+d^2)^{7/2}} \end{bmatrix}$$

Note that the $\frac{\partial^2 u}{\partial P \partial d} = \frac{\partial^2 u}{\partial d \partial P}$ given that the function $u(P, d)$ is smooth enough.

(d) 4 points

No, since the two chambers may have different locations(x, y, d) and different P.

The model parameter vector for two magma chambers should be $[P_1, P_2, d_1, d_2, x_1, x_2, y_1, y_2]^T$ and the observed vertical displacement should be:

$u = \frac{P_1 d_1}{((x-x_1)^2+(y-y_1)^2+d_1^2)^{3/2}} + \frac{P_2 d_2}{((x-x_2)^2+(y-y_2)^2+d_2^2)^{3/2}}$

If there are two magma chambers , it is not easy to tell their locations (x, y) in the first place. So they need to be inverted from observations. Also we need to have two different depth (d) parameters and P(pressure related parameters) to describe the two magma chamber.

There are some other acceptable solutions that assume the two magma chambers are at the same location and/or have the same P ...

(e) 4 points

8(Or other acceptable solutions)

(f) 4 points

No, because there are 8 degrees of freedom in the model space while only 6 degrees of freedom in the data space.

To map from a 8 dimensional model space to 6 dimensional data space means that there will be different models being mapped to the same point in data space.

This causes the inversion to have multiple solutions and trade off between different parameters.

There are also other acceptable solutions. If your number of model parameters is smaller than 6, then the solution may not exist. There may not be a set of four parameters that fits the 6 data points exactly. This is due to error in both model and measurements.

2 Problem 2 (graded by Yiran) - 25 points

(a) 6 points

Let the matrix $\boldsymbol{A} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4]$, then

$$
\begin{aligned}
\boldsymbol{A} = & \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \\
\boldsymbol{A}^T\boldsymbol{A} = & \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 2 & 0 \\ 1 & 2 & 0 & 3 \end{bmatrix} \\
= & \begin{bmatrix} \boldsymbol{x}_1^T\boldsymbol{x}_1 & \boldsymbol{x}_1^T\boldsymbol{x}_2 & \boldsymbol{x}_1^T\boldsymbol{x}_3 & \boldsymbol{x}_1^T\boldsymbol{x}_4 \\ & \boldsymbol{x}_2^T\boldsymbol{x}_2 & \boldsymbol{x}_2^T\boldsymbol{x}_3 & \boldsymbol{x}_2^T\boldsymbol{x}_4 \\ & \text{sym.} & \boldsymbol{x}_3^T\boldsymbol{x}_3 & \boldsymbol{x}_3^T\boldsymbol{x}_4 \\ & & & \boldsymbol{x}_4^T\boldsymbol{x}_4 \end{bmatrix}
\end{aligned}
$$

Therefore, we see that $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$ are orthogonal to each other; and $\boldsymbol{x}_3$ and $\boldsymbol{x}_4$ are orthogonal.

(b) 6 points

There are two independent vectors in $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_4\}$, choosing any two from them, along with $\boldsymbol{x}_3$, which is independent with all of them, form a basis. All the bases are:

$$\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\},\ \{\boldsymbol{x}_1, \boldsymbol{x}_4, \boldsymbol{x}_3\},\ \{\boldsymbol{x}_2, \boldsymbol{x}_4, \boldsymbol{x}_3\}$$

(c) 8 points

By definition

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{X}\boldsymbol{D}$$

where

$$
\begin{aligned}
\boldsymbol{X} &= [\boldsymbol{x}_1 \quad \boldsymbol{x}_2 \quad \boldsymbol{x}_3] \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix} \\
\boldsymbol{D} &= \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \\
&= \begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{bmatrix}
\end{aligned}
$$

Then

$$\boldsymbol{A} = \boldsymbol{X}\boldsymbol{D}\boldsymbol{X^{-1}}$$

Since

$$
\begin{aligned}
det(\boldsymbol{X}) &= 2 \\
\boldsymbol{X}^{-1} &= 1/2 \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix}^T \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & -0.5 & 0.5 \end{bmatrix}
\end{aligned}
$$

Then

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & -0.5 & 0.5 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4.5 & 0.5 \\ 0 & 0.5 & 4.5 \end{bmatrix}$$

(d) 5 points

The vector is stretched by λ_1, λ_2, λ_3 along the corresponding eigenvector directions $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$.

Check it with $\boldsymbol{x} = [1 \quad 2 \quad 3]^T$.

Let $\boldsymbol{v}_1 = \boldsymbol{X^{-1}}\boldsymbol{x} = [1 \quad 2.5 \quad 0.5]^T$

$$
\begin{aligned}
\boldsymbol{x} = &\quad \boldsymbol{X}\boldsymbol{X}^{-1}\boldsymbol{x} \\
= &\quad \boldsymbol{X}\boldsymbol{v_1} \\
= &\quad [\boldsymbol{x}_1 \quad \boldsymbol{x}_2 \quad \boldsymbol{x}_3] \begin{bmatrix} 1 \\ 2.5 \\ 0.5 \end{bmatrix} \\
= &\quad 1 \cdot \boldsymbol{x}_1 + 2.5 \cdot \boldsymbol{x}_2 + 0.5 \cdot \boldsymbol{x}_3
\end{aligned}
$$

We see that $\boldsymbol{v}_1$ is the coordinate of $\boldsymbol{x}$ in the basis $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\}$.
Then

$$
\begin{aligned}
\boldsymbol{A}\boldsymbol{x} = &\quad \boldsymbol{X}\boldsymbol{D}\boldsymbol{X}^{-1}\boldsymbol{x} \\
= &\quad \boldsymbol{X}\boldsymbol{D}\boldsymbol{v}_1 \\
= &\quad [\lambda_1\boldsymbol{x}_1 \quad \lambda_2\boldsymbol{x}_2 \quad \lambda_3\boldsymbol{x}_3]\boldsymbol{v}_1 \\
= &\quad \lambda_1 \cdot 1 \cdot \boldsymbol{x}_1 + \lambda_2 \cdot 2.5 \cdot \boldsymbol{x}_2 + \lambda_3 \cdot 0.5 \cdot \boldsymbol{x}_3
\end{aligned}
$$

which we see is a stretch of $\boldsymbol{x}$ by λ_i in $\boldsymbol{x}_i$.
This transformation gives
$2 \cdot [1 \quad 0 \quad 0]^T + 12.5 \cdot [0 \quad 1 \quad 1]^T + 2 \cdot [0 \quad -1 \quad 1]^T = [2 \quad 10.5 \quad 14.5]^T$, which is equal to the direct multiplication of $\boldsymbol{A}\boldsymbol{x}$.

3 Problem 3 (graded by Kangchen) - 24+1 points

(a) 6 points

This implies $\boldsymbol{v}^T\boldsymbol{x} = 0$ since $\boldsymbol{v}$ in V means that $\boldsymbol{v}$ is orthogonal to $\boldsymbol{x}$.

(b) 6 points

This is explaining that the set V is a subspace of $\mathbf{R}^n$

$(a_1\boldsymbol{v_1} + a_2\boldsymbol{v}_2)^T x = a_1\boldsymbol{v}_1^T\boldsymbol{x} + a_2\boldsymbol{v}_2^T\boldsymbol{x} = a_1(0) + a_2(0) = 0$

The first equality is by the distributive axiom of matrix multiplication.

This implies that $(a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2)$ is also orthogonal to $\boldsymbol{x}$.

so $(a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2)$ is in V

(c) 6 points

Since $\boldsymbol{x}$ is an eigenvector of $\boldsymbol{A}$, we assume $\boldsymbol{A}\boldsymbol{x} = k\boldsymbol{x}$

$(\boldsymbol{A}\boldsymbol{x})^T x = \boldsymbol{v}^T\boldsymbol{A}^T\boldsymbol{x} = \boldsymbol{v}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{v}^T(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{v}^T k\boldsymbol{x} = k\boldsymbol{v}^T\boldsymbol{x} = 0$

So $\boldsymbol{A}\boldsymbol{v}$ is also orthogonal to $\boldsymbol{x}$,for all $\boldsymbol{v}$ in V

(d) 6 points

Given condition: For all $\boldsymbol{v}$ in V (V is not a $\mathbf{0}$ space)if $\boldsymbol{A}\boldsymbol{v}$is also in V ,then one can find one eigenvector of $\boldsymbol{A}$ in subspace V.

$\boldsymbol{A}$ is a $n \times n$ matrix.

First we can find one eigenvector $\mathbf{x_1}$ in $\mathbf{R}^n$and we get the subspace V_1, a subset of R^n that is orthogonal to $\boldsymbol{x}_1$.

Then we can find another eigenvector $\boldsymbol{x_2}$in V_1 and get the subspace V_2, a subset of V_1 that is orthogonal to $\boldsymbol{x}_1$and $\boldsymbol{x}_2$.

....

Then we find $\boldsymbol{x_{n-1}}$ in V_{n-2} and get the subspace V_{n-1}a subset of V_{n-2} that is orthogonal to $\boldsymbol{x_1}, \boldsymbol{x_2}, ...\boldsymbol{x_{n-1}}$.

V_{n-1} will be a one dimensional subspace and we can find the last eigenvector $\boldsymbol{x}_n$

Thus, we have found that all the eigenvectors are orthogonal to each other.

This proof is enough for our purpose. We also attach the following proof that is more strict in a mathematical sense. This is not necessary to understand.

1) one can find one eigenvector $\mathbf{x}$ in $\mathbf{R^n}$. Since $\mathbf{R}^n$satisfy Given condition. This is trivial.

2) if a subspace V satisfies Given condition, and we have an eigenvector $\boldsymbol{x}$ in V. Then we prove that the set U that contains all vectors in V that are orthogonal to $\boldsymbol{x}$ also satisfies Given condition.

It is shown in (b) that U forms a subspace .

for all$\boldsymbol{u}$ in U,$\boldsymbol{Au}$ is in V since U is a subset of V

we can write $\boldsymbol{Au} = c\boldsymbol{x} + \boldsymbol{w}, \boldsymbol{w}$is in U, multiply both sides by $\mathbf{x}^T$:

$\boldsymbol{x}^T\boldsymbol{Au} = c\boldsymbol{x}^T\boldsymbol{x} + \boldsymbol{x}^T\boldsymbol{w} = c\boldsymbol{x}^T\boldsymbol{x}$

note that on the left hand side:$\boldsymbol{x}^T\boldsymbol{Au} = \boldsymbol{u}^T\boldsymbol{A}^T\boldsymbol{x} = \boldsymbol{u}^T\boldsymbol{Ax} = k\boldsymbol{u}^T\boldsymbol{x} = 0$

So $c\boldsymbol{x}^T\boldsymbol{x} = c|\mathbf{x}|^2 = 0$,which implies that $c = 0$,

So for all $\boldsymbol{u}$ in U ,$\boldsymbol{Au}$ is also in U. So U satisfies Given condition.

3) That means we can always find oneeigen vector while the remaining subspace still satisfies Given condition until it shrinks to $\{\mathbf{0}\}$. Since each time the dimension of the space is decreased by 1, we can find n orthogonal eigen vectors.

4 Problem 4 (graded by Yiran) - 25 points

(See atttached Matlab code and figures)
Fig. 1 shows the function and the four roots.
Fig. 2 shows the root found from each initial guess x_0. The red lines mark the four roots. We also plot the function $f(x)$ in black.
Fig. 3 shows that for all the initial guesses, the method converges within a certain number of iterations.

There are four roots of this function. If the initial guess is close enough to a given root, Newton's method will generally find this root that is closest to the initial guess, unless the initial guess is close to an extreme point.
If the initial guess is far away from the roots or close to an extreme point, there is no guarantee that Newton's method will find the closest root. The jump in x is controlled by the gradient at the current point. If the current point is close to an extreme point (and note that there are many extreme points of this function), the small gradient will cause a large jump to the neighbor of another root.
To find all the roots of the function, we need to try different initial guesses.