

# 1 Problem 1

**a**

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

**b**

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

**c**

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

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**

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.

**e**

8

**f**

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.

# 2 Problem 2

**a**

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

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

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

**b**

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

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

**c**

By definition

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

where

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

Then

$$\mathbf{A} = \mathbf{XDX}^{-1}$$

Since

$$\det(\mathbf{X}) = 2$$

$$\begin{aligned}\mathbf{X}^{-1} &= 1/2 \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix}^T \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & -0.5 & 0.5 \end{bmatrix}\end{aligned}$$

Then

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

**d**

The vector is stretched by  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$  along the corresponding eigenvector directions  $\mathbf{x}_1$ ,  $\mathbf{x}_2$ ,  $\mathbf{x}_3$ .

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

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

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

We see that  $\mathbf{v}_1$  contains the coordinates of  $\mathbf{x}$  in the basis  $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ .

Then

$$\begin{aligned}\mathbf{Ax} &= \mathbf{XDX}^{-1}\mathbf{x} \\ &= \mathbf{XD}\mathbf{v}_1 \\ &= [\lambda_1\mathbf{x}_1 \ \lambda_2\mathbf{x}_2 \ \lambda_3\mathbf{x}_3]\mathbf{v}_1 \\ &= \lambda_1 \cdot 1 \cdot \mathbf{x}_1 + \lambda_2 \cdot 2.5 \cdot \mathbf{x}_2 + \lambda_3 \cdot 0.5 \cdot \mathbf{x}_3\end{aligned}$$

which we see is a stretch of  $\mathbf{x}$  by  $\lambda_i$  in  $\mathbf{x}_i$ .

This transformation gives

$2 \cdot [1 \ 0 \ 0]^T + 12.5 \cdot [0 \ 1 \ 1]^T + 2 \cdot [0 \ -1 \ 1]^T$   
 $= [2 \ 10.5 \ 14.5]^T$ , which is equal to the direct multiplication of  $\mathbf{Ax}$ .

### 3 Problem 3

**a**

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

**b**

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

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

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

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

**c**

Since  $\mathbf{x}$  is an eigenvector of  $\mathbf{A}$ , we assume  $\mathbf{A}\mathbf{x} = k\mathbf{x}$  where  $k \in \mathbf{R}$

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

So  $\mathbf{A}\mathbf{v}$  is also orthogonal to  $\mathbf{x} \forall \mathbf{v} \in V$

**d**

Given condition:  $\forall \mathbf{v} \in V (V \neq \{0\})$ , if  $\mathbf{A}\mathbf{v} \in V$  then one can find one eigenvector of  $\mathbf{A}$  in subspace  $V$ .

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

First we can find one eigenvector  $\mathbf{x}_1$  in  $\mathbf{R}^n$  and we get the subspace  $V_1 \subseteq \mathbf{R}^n$  that is orthogonal to  $\mathbf{x}_1$ .

Then we can find another eigenvector  $\mathbf{x}_2$  in  $V_1$  and get the subspace  $V_2 \subseteq V_1$  that is orthogonal to  $\mathbf{x}_1$  and  $\mathbf{x}_2$ .

.....

Finally we find  $\mathbf{x}_n$  in  $V_{n-1}$  and get the subspace  $V_n \subseteq V_{n-1}$  that is orthogonal to  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ .

To this step, we have found all the eigenvectors that is orthogonal to each other.

This proof is enough for our purpose. We also attached the following proof that is more strict in a mathematical sense.

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

2) if a subspace  $V$  satisfy Given condition, and we have an eigen vector  $\mathbf{x} \in V$ . Then we prove that the set  $U = \{\mathbf{u} | \mathbf{u}^T \mathbf{x} = 0, \mathbf{u} \in V\}$  also satisfy Given condition.

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

$\forall \mathbf{u} \in U, \mathbf{A}\mathbf{u} \in U$  since  $U \subseteq V$

we can write  $\mathbf{A}\mathbf{u} = c\mathbf{x} + \mathbf{w}, \mathbf{w} \in U, c \in \mathbf{R}$ , multiply both sides by  $\mathbf{x}^T$ :

$$\mathbf{x}^T \mathbf{A}\mathbf{u} = c\mathbf{x}^T \mathbf{x} + \mathbf{x}^T \mathbf{w} = c\mathbf{x}^T \mathbf{x}$$

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

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

So  $\mathbf{A}\mathbf{u} = \mathbf{w} \in U, \forall \mathbf{u} \in U$ . So  $U$  satisfy Given condition.

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

## 4 Problem 4

(See attached 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 dots mark the positions of (root, root). We see for  $x_0$  close to a certain root, the method will find this root. For  $x_0$  that is not so close to any of the roots, there is no guarantee that the method will find the root that is closest to the initial guess.

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 find this root that is closest to the initial guess.

Otherwise, if the initial guess is far away, there is no guarantee that Newton's method will find the closest root. The "jump" in  $x$  is controlled by the gradient at 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.