

CS205 Homework #2 Solutions

Problem 1

[Heath 3.29, page 152] Let \mathbf{v} be a nonzero n -vector. The hyperplane normal to \mathbf{v} is the $(n-1)$ -dimensional subspace of all vectors \mathbf{z} such that $\mathbf{v}^T \mathbf{z} = 0$. A *reflector* is a linear transformation \mathbf{R} such that $\mathbf{R}\mathbf{x} = -\mathbf{x}$ if \mathbf{x} is a scalar multiple of \mathbf{v} , and $\mathbf{R}\mathbf{x} = \mathbf{x}$ if $\mathbf{v}^T \mathbf{x} = 0$. Thus, the hyperplane acts as a mirror: for any vector, its component within the hyperplane is invariant, whereas its component orthogonal to the hyperplane is reversed.

1. Show that $\mathbf{R} = 2\mathbf{P} - \mathbf{I}$, where \mathbf{P} is the orthogonal projector onto the hyperplane normal to \mathbf{v} . Draw a picture to illustrate this result
2. Show that \mathbf{R} is symmetric and orthogonal
3. Show that the Householder transformation

$$\mathbf{H} = \mathbf{I} - 2 \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T \mathbf{v}},$$

is a reflector

4. Show that for any two vectors \mathbf{s} and \mathbf{t} such that $\mathbf{s} \neq \mathbf{t}$ and $\|\mathbf{s}\|_2 = \|\mathbf{t}\|_2$, there is a reflector \mathbf{R} such that $\mathbf{R}\mathbf{s} = \mathbf{t}$

Solution

1. We can obtain the reflection $\mathbf{R}\mathbf{x}$ of a vector \mathbf{x} with respect to a hyperplane through the origin by adding to \mathbf{x} twice the vector from \mathbf{x} to $\mathbf{P}\mathbf{x}$, where $\mathbf{P}\mathbf{x}$ is the projection of \mathbf{x} onto the same hyperplane (see figure 1). Thus

$$\mathbf{R}\mathbf{x} = \mathbf{x} + 2(\mathbf{P}\mathbf{x} - \mathbf{x}) = (2\mathbf{P}\mathbf{x} - \mathbf{x}) = (2\mathbf{P} - \mathbf{I})\mathbf{x}$$

Since this has to hold for all \mathbf{x} we have $\mathbf{R} = 2\mathbf{P} - \mathbf{I}$.

An alternative way to derive the same result is to observe that the projection $\mathbf{P}\mathbf{x}$ lies halfway between \mathbf{x} and its reflection $\mathbf{R}\mathbf{x}$. Therefore

$$\frac{1}{2}(\mathbf{x} + \mathbf{R}\mathbf{x}) = \mathbf{P}\mathbf{x} \Rightarrow \mathbf{R}\mathbf{x} = (2\mathbf{P}\mathbf{x} - \mathbf{x}) = (2\mathbf{P} - \mathbf{I})\mathbf{x}$$

which leads to the same result.

2. A reflection with respect to a hyperplane through the origin does not change the magnitude of the reflected vector (see figure 1). Therefore we have

$$\|\mathbf{R}\mathbf{x}\| = \|\mathbf{x}\| \Rightarrow \mathbf{x}^T \mathbf{R}^T \mathbf{R} \mathbf{x} = \mathbf{x}^T \mathbf{x} \Rightarrow \mathbf{x}^T (\mathbf{R}^T \mathbf{R} - \mathbf{I}) \mathbf{x} = 0$$

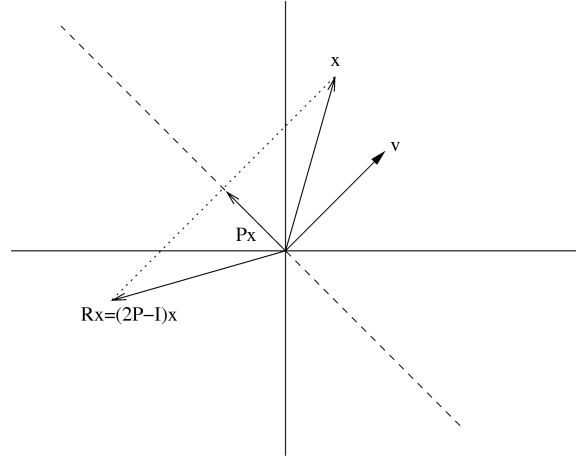


Figure 1: Reflector

for any vector \mathbf{x} . If we could show that $\mathbf{x}^T(\mathbf{R}^T\mathbf{R} - \mathbf{I})\mathbf{x} = 0$ implies $\mathbf{R}^T\mathbf{R} - \mathbf{I} = \mathbf{0}$ we would have proven the orthogonality of \mathbf{R} . Furthermore, since reflecting a vector twice just gives the original vector we have $\mathbf{R}^2 = \mathbf{I}$. Therefore we would have

$$\mathbf{R}^T\mathbf{R} = \mathbf{I} \Rightarrow \mathbf{R}^T\mathbf{R}^2 = \mathbf{R} \Rightarrow \mathbf{R}^T = \mathbf{R}$$

which shows that \mathbf{R} is symmetric.

In order to show that $\mathbf{R}^T\mathbf{R} - \mathbf{I} = \mathbf{0}$ it suffices to show that for a symmetric matrix \mathbf{C} , $\mathbf{x}^T\mathbf{C}\mathbf{x} = 0$ for all \mathbf{x} implies $\mathbf{C} = \mathbf{0}$ (since $\mathbf{R}^T\mathbf{R} - \mathbf{I}$ is symmetric). To show that, we note that $\mathbf{e}_i^T\mathbf{C}\mathbf{e}_j = C_{ij}$ where \mathbf{e}_k is the k -th column of the identity matrix. We have $\mathbf{e}_i^T\mathbf{C}\mathbf{e}_i = C_{ii} = 0$ for any i and furthermore

$$0 = (\mathbf{e}_i + \mathbf{e}_j)^T \mathbf{C} (\mathbf{e}_i + \mathbf{e}_j) = C_{ii} + C_{jj} + C_{ij} + C_{ji} = 2C_{ij} = 2C_{ji} \Rightarrow \mathbf{C} = \mathbf{0}$$

3. The Householder matrix reflects all vectors in the direction of \mathbf{v}

$$\mathbf{H}(\alpha\mathbf{v}) = \left(\mathbf{I} - 2\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}} \right) (\alpha\mathbf{v}) = \alpha\mathbf{v} - 2\alpha\frac{\mathbf{v}(\mathbf{v}^T\mathbf{v})}{\mathbf{v}^T\mathbf{v}} = \alpha(\mathbf{v} - 2\mathbf{v}) = -(\alpha\mathbf{v})$$

and leaves all vectors \mathbf{x} with $\mathbf{v}^T\mathbf{x} = 0$ invariant

$$\mathbf{H}\mathbf{x} = \left(\mathbf{I} - 2\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}} \right) \mathbf{x} = \mathbf{x} - 2\frac{\mathbf{v}(\mathbf{v}^T\mathbf{x})}{\mathbf{v}^T\mathbf{v}} = \mathbf{x}$$

therefore, \mathbf{H} is a reflector about the hyperplane $\{x : \mathbf{v}^T\mathbf{x} = 0\}$.

4. Any two vectors \mathbf{s} and \mathbf{t} are reflections of each other with respect to the hyperplane normal to the vector $\mathbf{s} - \mathbf{t}$ that passes from the midpoint of \mathbf{s} and \mathbf{t} . When $\|\mathbf{s}\|_2 = \|\mathbf{t}\|_2$

that hyperplane passes through the origin and can be written as $\{x : (s - t)^T x = 0\}$. Therefore the Householder transform

$$\mathbf{H} = \mathbf{I} - 2 \frac{(s - t)(s - t)^T}{(s - t)^T(s - t)}$$

is the reflection that maps s to t and vice versa.

To show that formally, we have

$$\begin{aligned} \mathbf{H}s &= s - 2 \frac{(s - t)(s - t)^T s}{(s - t)^T(s - t)} = \frac{s(s - t)^T(s - t) - 2(s - t)(s - t)^T s}{(s - t)^T(s - t)} \\ &= \frac{ss^T s - 2st^T s + st^T t - 2ss^T s + 2st^T s + 2ts^T s - 2tt^T s}{(s - t)^T(s - t)} \\ &= \frac{2ts^T s - 2tt^T s}{(s - t)^T(s - t)} \begin{bmatrix} ss^T s + st^T t - 2ss^T s = 0 \\ -2st^T s + 2st^T s = 0 \end{bmatrix} \\ &= \frac{ts^T s - 2tt^T s + tt^T t}{(s - t)^T(s - t)} = \frac{t(s - t)^T(s - t)}{(s - t)^T(s - t)} = t \end{aligned}$$

Problem 2

Let \mathbf{A} be a rectangular $m \times n$ matrix with full column rank and $m > n$. Consider the **QR** decomposition of \mathbf{A} .

1. Show that $\mathbf{P}_0 = \mathbf{I} - \mathbf{Q}\mathbf{Q}^T$ is the projection matrix onto the nullspace of \mathbf{A}^T
2. Show that for every \mathbf{x} we have $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 = \|\mathbf{A}(\mathbf{x} - \mathbf{x}_0)\|_2^2 + \|\mathbf{A}\mathbf{x}_0 - \mathbf{b}\|_2^2$ where \mathbf{x}_0 is the least squares solution of $\mathbf{A}\mathbf{x} = \mathbf{b}$
3. Show that the minimum value for the 2-norm of the residual is attained when \mathbf{x} is equal to the least squares solution and that this minimum value is equal to $\|\mathbf{P}_0\mathbf{b}\|_2$

Solution

1. We know that the nullspace of \mathbf{A}^T and the column space of \mathbf{A} are the normal complements of each other. Therefore, any vector \mathbf{x} can be written as $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ where \mathbf{x}_1 is in the nullspace of \mathbf{A}^T and \mathbf{x}_2 is in the column space of \mathbf{A} .

For \mathbf{x}_1 this means that $\mathbf{A}^T \mathbf{x}_1 = \mathbf{0}$. Since \mathbf{A} is full rank, the **QR** decomposition is defined and

$$\mathbf{A}^T \mathbf{x}_1 = \mathbf{0} \Rightarrow \mathbf{R}^T \mathbf{Q}^T \mathbf{x}_1 = \mathbf{0} \Rightarrow \mathbf{Q}^T \mathbf{x}_1 = \mathbf{0}$$

since \mathbf{R} is nonsingular. On the other hand, \mathbf{x}_2 belongs to the column space of \mathbf{A} , therefore it can be written as $\mathbf{x}_2 = \mathbf{A}\mathbf{y} = \mathbf{Q}\mathbf{R}\mathbf{y}$ where $\mathbf{y} \in \mathbb{R}^n$.

Thus, the action of \mathbf{P}_0 on \mathbf{x} amounts to

$$\begin{aligned}\mathbf{P}_0\mathbf{x} &= (\mathbf{I} - \mathbf{Q}\mathbf{Q}^T)(\mathbf{x}_1 + \mathbf{x}_2) = \mathbf{x}_1 + \mathbf{x}_2 - \mathbf{Q}\mathbf{Q}^T\mathbf{x}_1 - \mathbf{Q}\mathbf{Q}^T\mathbf{Q}\mathbf{R}\mathbf{y} \\ &= \mathbf{x}_1 + \mathbf{x}_2 - \mathbf{Q}\mathbf{R}\mathbf{y} = \mathbf{x}_1\end{aligned}$$

Therefore, \mathbf{P}_0 is the projection matrix onto the nullspace of \mathbf{A}^T (and $\mathbf{Q}\mathbf{Q}^T$ is the projection matrix onto the column space of \mathbf{A}).

2. We have

$$\begin{aligned}\|\mathbf{Ax} - \mathbf{b}\|_2^2 &= \|\mathbf{A}(\mathbf{x} - \mathbf{x}_0) + (\mathbf{Ax}_0 - \mathbf{b})\|_2^2 \\ &= [\mathbf{A}(\mathbf{x} - \mathbf{x}_0) + (\mathbf{Ax}_0 - \mathbf{b})]^T [\mathbf{A}(\mathbf{x} - \mathbf{x}_0) + (\mathbf{Ax}_0 - \mathbf{b})] \\ &= \|\mathbf{A}(\mathbf{x} - \mathbf{x}_0)\|_2^2 + \|\mathbf{Ax}_0 - \mathbf{b}\|_2^2 + 2(\mathbf{x} - \mathbf{x}_0)^T \mathbf{A}^T (\mathbf{Ax}_0 - \mathbf{b}) \\ &= \|\mathbf{A}(\mathbf{x} - \mathbf{x}_0)\|_2^2 + \|\mathbf{Ax}_0 - \mathbf{b}\|_2^2 + 2(\mathbf{x} - \mathbf{x}_0)^T (\mathbf{A}^T \mathbf{Ax}_0 - \mathbf{A}^T \mathbf{b}) \\ &= \|\mathbf{A}(\mathbf{x} - \mathbf{x}_0)\|_2^2 + \|\mathbf{Ax}_0 - \mathbf{b}\|_2^2 \quad [\mathbf{A}^T \mathbf{Ax}_0 - \mathbf{A}^T \mathbf{b} = \mathbf{0}]\end{aligned}$$

3. From the equation above, we have that the minimum value for $\|\mathbf{Ax} - \mathbf{b}\|_2$ is attained for $\mathbf{x} = \mathbf{x}_0$, since the term $\|\mathbf{Ax}_0 - \mathbf{b}\|_2^2$ does not depend on \mathbf{x} . The least squares solution is given as

$$\mathbf{Rx}_0 = \mathbf{Q}^T \mathbf{b} \Rightarrow \mathbf{x}_0 = \mathbf{R}^{-1} \mathbf{Q}^T \mathbf{b}$$

Therefore the minimum value for the residual $\|\mathbf{Ax} - \mathbf{b}\|_2$ is

$$\begin{aligned}\|\mathbf{Ax}_0 - \mathbf{b}\|_2 &= \|\mathbf{Q}\mathbf{R}\mathbf{R}^{-1}\mathbf{Q}^T \mathbf{b} - \mathbf{b}\|_2 = \|\mathbf{Q}\mathbf{Q}^T \mathbf{b} - \mathbf{b}\|_2 = \|(\mathbf{Q}\mathbf{Q}^T - \mathbf{I})\mathbf{b}\|_2 \\ &= \|(\mathbf{I} - \mathbf{Q}\mathbf{Q}^T)\mathbf{b}\|_2 = \|\mathbf{P}_0 \mathbf{b}\|_2\end{aligned}$$

Intuitively, this means that the least squares solution annihilates the component of the residual in the column space of \mathbf{A} and the minimum value for the residual is exactly the component of \mathbf{b} that is *not* contained in the column space of \mathbf{A} .

Problem 3

State whether the following classes of matrices are positive (semi-)definite, negative (semi-)definite, indefinite, or whether their definiteness cannot be determined in general

1. Orthogonal matrices
2. Matrices of the form $\mathbf{A}^T \mathbf{A}$ where \mathbf{A} is a rectangular matrix
3. Projection matrices
4. Matrices of the form $\mathbf{I} - \mathbf{P}$ where \mathbf{P} is a projection matrix
5. Householder matrices

6. Upper triangular matrices with positive diagonal elements
7. A diagonally dominant matrix with positive elements on the diagonal. A matrix is called diagonally dominant if $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ and $|a_{ii}| > \sum_{j \neq i} |a_{ji}|$.

Solution

1. Any diagonal matrix with values $+1$ or -1 on the diagonal is orthogonal. Nevertheless it can be positive definite (if it equals \mathbf{I}), negative definite (if it equals $-\mathbf{I}$) or indefinite (in any other case). Thus the definiteness of orthogonal matrices cannot be determined for the general case.
2. We have $\mathbf{x}^T(\mathbf{A}^T\mathbf{A})\mathbf{x} = \|\mathbf{Ax}\|_2^2 \geq 0$. Thus a matrix of the form $\mathbf{A}^T\mathbf{A}$ is always positive semidefinite. In addition, if \mathbf{A} is full rank, then $\mathbf{A}^T\mathbf{A}$ is positive definite (since $\mathbf{Ax} = \mathbf{0} \Rightarrow \mathbf{x} = \mathbf{0}$).
3. Let V be the vector subspace that a projection matrix \mathbf{P} projects onto, and V^\perp its normal complement. Let $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ be an arbitrary vector, where \mathbf{x}_1 is the component of \mathbf{x} in V and \mathbf{x}_2 its component in V^\perp . Therefore

$$\mathbf{x}^T\mathbf{Px} = (\mathbf{x}_1 + \mathbf{x}_2)^T\mathbf{P}(\mathbf{x}_1 + \mathbf{x}_2) = \mathbf{x}_1^T\mathbf{Px}_1 = \|\mathbf{x}_1\|_2^2 \geq 0$$

where we used the fact that \mathbf{P} is symmetric and $\mathbf{Px}_2 = \mathbf{0}$. Therefore a projection matrix is always positive semi-definite.

4. The matrix $\mathbf{I} - \mathbf{P}$ is the projection onto the *normal complement* of the space \mathbf{P} projects onto. Therefore it is a projection matrix itself and thus positive semidefinite.
5. Given the Householder matrix

$$\mathbf{H} = \mathbf{I} - 2 \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}$$

we have $\mathbf{v}^T\mathbf{H}\mathbf{v} = \mathbf{v}^T(-\mathbf{v}) = -\|\mathbf{v}\|_2^2 < 0$ where if \mathbf{w} is a nonzero vector that is orthogonal to \mathbf{v} (such a vector always exists in 2 or more dimensions) then $\mathbf{w}^T\mathbf{H}\mathbf{w} = \mathbf{w}^T\mathbf{w} = \|\mathbf{w}\|_2^2 > 0$. Therefore a Householder matrix is always indefinite (in the special 1D case the matrix reduces to the single number -1 , being negative definite)

6. \mathbf{I} is an example of such a matrix that is positive definite. The matrix

$$\begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix}$$

is an example of an indefinite matrix. However, we can conclude that such a matrix can never be negative definite, because $\mathbf{e}_i^T\mathbf{A}\mathbf{e}_i = \mathbf{A}_{ii} > 0$, where \mathbf{e}_i is the i -th column of the identity matrix.

7.

$$\begin{aligned}
\mathbf{x}^T \mathbf{A} \mathbf{x} &= \sum_{i,j} \mathbf{A}_{ij} x_i x_j = \sum_i \mathbf{A}_{ii} x_i^2 + \sum_{i,j \neq i} \mathbf{A}_{ij} x_i x_j \\
&\geq \sum_i |\mathbf{A}_{ii}| |x_i|^2 - \sum_{i,j \neq i} |\mathbf{A}_{ij}| |x_i| |x_j| \\
&= \frac{1}{2} \sum_i (|\mathbf{A}_{ii}| + |\mathbf{A}_{ii}|) |x_i|^2 - \sum_{i,j \neq i} |\mathbf{A}_{ij}| |x_i| |x_j| \\
&> \frac{1}{2} \sum_{i,j \neq i} (|\mathbf{A}_{ij}| + |\mathbf{A}_{ji}|) |x_i|^2 - \sum_{i,j \neq i} |\mathbf{A}_{ij}| |x_i| |x_j| \\
&= \sum_{i,j \neq i} |\mathbf{A}_{ij}| \left(\frac{1}{2} |x_i|^2 + \frac{1}{2} |x_j|^2 - |x_i| |x_j| \right) \\
&= \frac{1}{2} \sum_{i,j \neq i} |\mathbf{A}_{ij}| (|x_i| - |x_j|)^2 \geq 0
\end{aligned}$$

Problem 4

[Heath 3.12 page 150]

- Let \mathbf{A} be a $n \times n$ matrix. Show that any two of the following conditions imply the other.
 - $\mathbf{A}^T = \mathbf{A}$
 - $\mathbf{A}^T \mathbf{A} = \mathbf{I}$
 - $\mathbf{A}^2 = \mathbf{I}$
- Give a specific example, other than the identity matrix \mathbf{I} or a permutation of it, of a 3×3 matrix that has all three of these properties.
- Name a nontrivial class of matrices that have all three of these properties.

Solution

1.

$$\left\{ \begin{array}{l} \mathbf{A}^T = \mathbf{A} \\ \mathbf{A}^T \mathbf{A} = \mathbf{I} \end{array} \right\} \Rightarrow \mathbf{A}^2 = \mathbf{I} \quad [\text{By substitution}]$$

$$\left\{ \begin{array}{l} \mathbf{A}^T = \mathbf{A} \\ \mathbf{A}^2 = \mathbf{I} \end{array} \right\} \Rightarrow \mathbf{A}^T \mathbf{A} = \mathbf{I} \quad [\text{By substitution}]$$

$$\left\{ \begin{array}{l} \mathbf{A}^T \mathbf{A} = \mathbf{I} \\ \mathbf{A}^2 = \mathbf{I} \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \mathbf{A}^{-1} = \mathbf{A}^T \\ \mathbf{A}^{-1} = \mathbf{A} \end{array} \right\} \Rightarrow \mathbf{A}^T = \mathbf{A}$$

2.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \begin{bmatrix} 1/3 & -2/3 & -2/3 \\ -2/3 & 1/3 & -2/3 \\ -2/3 & -2/3 & 1/3 \end{bmatrix}$$

3. Reflection matrices (e.g. Householder matrices, see problem 1).

Problem 5

[Heath 3.16 page 150] Consider the vector \mathbf{a} as an $n \times 1$ matrix.

1. Write out its QR factorization, showing the matrices \mathbf{Q} and \mathbf{R} explicitly.
2. What is the solution to the linear least squared problem $\mathbf{a}\mathbf{x} \cong \mathbf{b}$, where \mathbf{b} is a given n -vector?

Solution

1. By simple application of the algorithm, we have

$$\mathbf{Q} = \frac{1}{\|\mathbf{a}\|_2} \mathbf{a} \quad \mathbf{R} = [\|\mathbf{a}\|_2]$$

2. The least squares solution is given by the equation

$$\mathbf{R}x = \mathbf{Q}^T \mathbf{b} \Rightarrow \|\mathbf{a}\|_2 \cdot x = \frac{1}{\|\mathbf{a}\|_2} \mathbf{a}^T \mathbf{b} \Rightarrow x = \frac{\mathbf{a}^T \mathbf{b}}{\mathbf{a}^T \mathbf{a}}$$