Georgia Institute of Technology

ECE 6555 - Assignment 1

due Thursday September 8, 2022 - v1.2

- There are 8 problems over 7 pages (including the cover page).
- The problems are not necessarily in order of difficulty.
- Every question in a problem is worth 2 points, so problems with many questions are worth more than problems with few questions.
- Each question is graded as follows: no credit without meaningful work, half credit for partial work, full credit if essentially correct.
- Unless otherwise specified, you should concisely indicate your reasoning and show all relevant work.
- The grade on each question is based on our judgment of your level of understanding as reflected by what you have written. If we cannot read it, we cannot grade it.
- Please use a pen and not a pencil if you handwrite your solution.
- **You must submit your assignment on Gradescope**.

Problem 1: Space of solutions

Let $\mathbf{y} \in \mathbb{R}^m$ and $\mathbf{H} \in \mathbb{R}^{m\times n}$. Consider the equation $\mathbf{Hx} = \mathbf{y}$ and assume that $\mathbf{y} \in \mathrm{Im}(\mathbf{H})$ so that there exists at least one solution $\mathbf{x}_0$. Show that

$$\{\mathbf{x} : \mathbf{Hx} = \mathbf{y}\} = \mathbf{x}_0 + \mathrm{Ker}(\mathbf{H})$$

SOLUTION Since $\mathbf{y} \in \mathrm{Im}(\mathbf{H})$, there exists $\mathbf{x}_0$ such that $\mathbf{y} = \mathbf{Hx}_0$. Hence the equation $\mathbf{Hx} = \mathbf{y}$ admits at least $\mathbf{x}_0$ as a solution.

Let $\mathbf{x} \in \mathbf{x}_0 + \mathrm{Ker}(\mathbf{H})$. Then $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_1$ where $\mathbf{x}_1$ is such that $\mathbf{Hx}_1 = 0$. Hence, $\mathbf{Hx} = \mathbf{Hx}_0 + \mathbf{Hx}_1 = \mathbf{y} + 0 = \mathbf{y}$ and $\mathbf{x}_0 + \mathrm{Ker}(\mathbf{H}) \subset \{\mathbf{x} : \mathbf{Hx} = \mathbf{y}\}$.

Let $\mathbf{x}^* \in \{\mathbf{x} : \mathbf{Hx} = \mathbf{y}\}$. Then we can write $\mathbf{x}^* = \mathbf{x}^* + \mathbf{x}_0 - \mathbf{x}_0 = \mathbf{x}_0 + \mathbf{x}_2$ where $\mathbf{x}_2 \triangleq \mathbf{x}^* - \mathbf{x}_0$. Note that $\mathbf{Hx}_2 = \mathbf{Hx}^* - \mathbf{Hx}_0 = \mathbf{y} - \mathbf{y} = 0$ so that $\mathbf{x}_2 \in \mathrm{Ker}(\mathbf{H})$. Hence $\{\mathbf{x} : \mathbf{Hx} = \mathbf{y}\} \subset \mathbf{x}_0 + \mathrm{Ker}(\mathbf{H})$. □

Problem 2: Inner product

[Q1] Consider a vector space V over the field $\mathbb{R}$ with inner product $\langle \cdot, \cdot, \rangle$. Show that $\|\mathbf{x}\| \triangleq \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ is indeed a norm.

SOLUTION We need to check the properties of a norm.

- *Triangle inequality:*

$$\langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle$$

Note that $\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle^*$ is real-valued and therefore by Cauchy-Schwartz

$$\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle \leq 2|\langle \mathbf{x}, \mathbf{y} \rangle| \leq 2\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}\sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}.$$

Hence,

$$\langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle \leq \left(\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} + \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}\right)^2$$

- *Absolute homogeneity* We have

$$\sqrt{\langle \lambda\mathbf{x}, \lambda\mathbf{x} \rangle} = \sqrt{|\lambda|^2 \langle \mathbf{x}, \mathbf{x} \rangle} = |\lambda|\sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

- *Positive definiteness* We have $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ with equality iff $= \mathbf{0}$.

□

[Q2] In this question, we will operate with the field of complex numbers $\mathbb{C}$ instead of $\mathbb{R}$. The operation $\dagger$ denotes the transpose conjugate, e.g., $\mathbf{x}^\dagger \triangleq (\mathbf{x}^*)^T$. An inner product must now be a symmetric sesquilinear form instead of symmetric linear form, i.e.,

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle^*$$

Show that the function $\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C} : (\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x}^T\mathbf{y}^* = \mathbf{y}^\dagger\mathbf{x}$ is a valid inner product. Conclude that $\|\mathbf{x}\| \triangleq \sqrt{\mathbf{x}^\dagger\mathbf{x}}$ is a norm.

SOLUTION We need to check the properties of an inner product.

- *Linearity:* $\mathbf{y}^\dagger(\alpha\mathbf{x}+\beta\mathbf{x}') = \alpha\mathbf{y}^\dagger\mathbf{x}+\beta\mathbf{y}^\dagger\mathbf{x}'$
- *Reflexivity:* $\mathbf{y}^\dagger\mathbf{x} = (\mathbf{x}^\dagger\mathbf{y})^*$
- *Non degenerate* $\mathbf{x}^\dagger\mathbf{x} = \sum_{i=1}^n |x_i|^2 \geq 0$ with equality iff $\mathbf{x}=0$.

□

Problem 3: Orthogonal complement

For a subvector space W of a vector space V (of finite dimension and over $\mathbb{R}$), the orthogonal complement is

$$W^\perp \triangleq \{v \in V : \forall w \in W\ \langle v, w\rangle = 0\}.$$

[Q1] Show that $W^\perp$ is a vector subspace.

SOLUTION Let $v_1, v_2 \in W^\perp$ and $\alpha, \beta \in \mathbb{C}$. Then for all $w \in W$

$$\langle \alpha v_1 + \beta v_2, w\rangle = \alpha\langle v_1, w\rangle + \beta\langle v_2, w\rangle = 0.$$

In addition $0 \in W^\perp$ so that $W^\perp$ is a subvector space of V. □

[Q2] Show that $W \oplus W^\perp = V$.

SOLUTION Note that if $v \in W \cap W^\perp$ then $\langle v, v\rangle = 0$ so that $v = 0$. Futhermore, if $v \neq 0 \in W$ and $w \neq 0 \in W^\perp$, then

$$\alpha v + \beta w = 0 \Rightarrow \begin{cases} \alpha\langle v, v\rangle + \beta\langle v, w\rangle = 0 \\ \alpha\langle v, w\rangle + \beta\langle w, w\rangle = 0 \end{cases} \Rightarrow \begin{cases} \alpha\|v\|^2 = 0 \\ \beta\|w\|^2 = 0 \end{cases} \Rightarrow \begin{cases} \alpha = \beta = 0 \end{cases}$$

so that v and w are independent. Hence, W and $W^\perp$ are in direct sum.

Finally, all that remains to show is that every vector $v \in V$ can be written as $v = w_1 + w_2$ where $w_1 \in W$ and $w_2 \in W^\perp$. Consider an orthonormal basis $\{e_i\}_{i=1}^k$ for W and consider the vector

$$w_1 = \sum_{i=1}^k \langle w, e_i\rangle e_i \in W.$$

and form the vector $w_2 = w - w_1 = w - \sum_{i=1}^k \langle w, e_i\rangle e_i$. Note that for every $j \in [1; k]$

$$\langle w_2, e_j\rangle = \langle w, e_j\rangle - \sum_{i=1}^k \langle w, e_i\rangle\langle e_i, e_j\rangle = 0.$$

Hence $w_2 \in W^\perp$. This also shows that $\dim(W) + \dim(W^\perp) = \dim(V)$. □

[Q3] Show that $(W^\perp)^\perp = W$.

SOLUTION Let $w \in W$. Then $\forall u \in W^\perp$, $\langle w, u\rangle = 0$, which means that $w \in (W^\perp)^\perp$. Hence $W \subset (W^\perp)^\perp$. In addition

$$\dim((W^\perp)^\perp) + \dim(W^\perp) = \dim(V) = \dim(W) + \dim(W^\perp) = \dim(V)$$

and we have $\dim(W) = \dim((W^\perp)^\perp)$ so that $W = (W^\perp)^\perp$. □

Problem 4: Row space

Let $\mathbf{H} \in \mathbb{R}^{m\times n}$. Show that $\text{Im}(\mathbf{H}^T\mathbf{H}) = \text{Im}(\mathbf{H}^T)$.

SOLUTION By definition, $\text{Im}(\mathbf{H}^T\mathbf{H}) \subset \text{Im}(\mathbf{H}^T)$. Now, for every $y \in \text{Im}(\mathbf{H}^T)$ there exists $\mathbf{x} \in \mathbb{R}^m$ such that $y = \mathbf{H}^T\mathbf{x}$. We proved in class that $\text{Im}(\mathbf{H}) + \text{Ker}(\mathbf{H}^\top) = \mathbb{R}^m$ so that we have $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ with $\mathbf{x}_1 \in \text{Im}(\mathbf{H})$ and $\mathbf{x}_2 \in \text{Ker}(\mathbf{H}^\top)$. Hence $\mathbf{y} = \mathbf{H}^T\mathbf{H}\mathbf{x}_1$ and $\mathbf{y} \in \text{Im}(\mathbf{H}^T\mathbf{H})$.

$\square$

Problem 5: Weighted least-squares

Consider a modified cost function

$$J(\mathbf{x}) \triangleq (\mathbf{y} - \mathbf{H}\mathbf{x})^T\mathbf{W}(\mathbf{y} - \mathbf{H}\mathbf{x})$$

where $\mathbf{W}$ is symmetric positive definite.

[Q1] Show that the cost is always positive.

SOLUTION This actually follows from the definition of being positive definite, which ensures that $\mathbf{x}^T\mathbf{W}\mathbf{x} > 0$ for all $\mathbf{x} \neq 0$. If you prefer to define definite positive definite in terms of eigenvalues, write $W = \mathbf{U}\mathbf{D}^{\frac{1}{2}}\mathbf{D}^{\frac{1}{2}}\mathbf{U}^T$ where $\mathbf{D}^{\frac{1}{2}}$ is a diagonal matrix with real-valued positive diagonal entries. Hence,

$$\begin{aligned} J(\mathbf{x}) &\triangleq (\mathbf{y} - \mathbf{H}\mathbf{x})^T\mathbf{W}(\mathbf{y} - \mathbf{H}\mathbf{x}) \\ &= (\mathbf{y} - \mathbf{H}\mathbf{x})^T\mathbf{U}\mathbf{D}^{\frac{1}{2}}\mathbf{D}^{\frac{1}{2}}\mathbf{U}^T(\mathbf{y} - \mathbf{H}\mathbf{x}) \\ &= \left\|\mathbf{D}^{\frac{1}{2}}\mathbf{U}^T(\mathbf{y} - \mathbf{H}\mathbf{x})\right\|_2^2 \geq 0. \end{aligned}$$

$\square$

[Q2] Show that the cost is zero if an only if $\mathbf{y} = \mathbf{H}\mathbf{x}$.

SOLUTION The cost is zero if and only if $\left\|\mathbf{D}^{\frac{1}{2}}\mathbf{U}^T(\mathbf{y} - \mathbf{H}\mathbf{x})\right\|_2 = 0$, i.e., if $\mathbf{y} - \mathbf{H}\mathbf{x}$ is in the kernel of $\mathbf{D}^{\frac{1}{2}}\mathbf{U}^T$. Because $\mathbf{U}$ is unitary and $\mathbf{D}$ has positive entries, this means that $\mathbf{y} = \mathbf{H}\mathbf{x}$. $\square$

[Q3] Show that the normal equations corresponding to that modified cost are $\mathbf{H}^T\mathbf{W}\mathbf{H}\mathbf{x} = \mathbf{H}^T\mathbf{W}\mathbf{y}$.

SOLUTION The normal equations were obtained by taking the gradient of the cost and writing stationary conditions. Here the gradient of $J(\mathbf{x})$ is (assuming everything is real-valued to make my life easier)

$$-2\mathbf{H}^T\mathbf{W}\mathbf{y} + 2\mathbf{H}^T\mathbf{W}\mathbf{H}\mathbf{x}$$

from which the solution follows. $\square$

Problem 6: Projectors

In class, we defined $P_{\mathrm{H}} \triangleq \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T$ as the projector on $\text{Im}(\mathbf{H})$.

[Q1] Verify that P_{H} is symmetric and idempotent (i.e., $P_{\mathrm{H}} = P_{\mathrm{H}}^T$).

SOLUTION

$$P_{\mathrm{H}}^T = (\mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T)^T = \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T = P_{\mathrm{H}}$$

$$P_{\mathrm{H}}^2 = \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\underbrace{\mathbf{H}^T\mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}}_{\mathrm{I}}\mathbf{H}^T = P_{\mathrm{H}}$$

$\square$

[Q2] Show that $\mathbf{I} - P_{\mathrm{H}}$ is a projector on $\mathrm{Im}(\mathbf{H})^{\perp}$.

SOLUTION Let $\mathbf{y} \in \mathbb{C}^n$ and let $\mathbf{z} = (\mathbf{I} - P_{\mathrm{H}})\mathbf{y}$. Then for all $\mathbf{x} \in \mathrm{Im}(\mathbf{H})$, there exists $\mathbf{u}$ such that $\mathbf{x} = \mathbf{Hu}$ and

$$\mathbf{x}^T\mathbf{z} = \mathbf{u}\mathbf{H}^T(\mathbf{I} - P_{\mathrm{H}})\mathbf{y} = \mathbf{u}\mathbf{H}^T\mathbf{y} - \mathbf{u}\mathbf{H}^T\mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{y} = 0.$$

Hence $\mathbf{z} \in \mathrm{Im}(\mathbf{H})^{\perp}$.

Now, for all $\mathbf{x} \in \mathrm{Im}(\mathbf{H})^{\perp}$

$$\mathbf{x}^T(\mathbf{z} - \mathbf{y}) = \mathbf{x}^T(-P_{\mathrm{H}}\mathbf{y}) = 0$$

so that the error is orthogonal to $\mathrm{Im}(\mathbf{H})^{\perp}$. Hence $\mathbf{I} - P_{\mathrm{H}}$ is indeed the orthogonal projection on $\mathrm{Im}(\mathbf{H})^{\perp}$. $\square$

[Q3] Show that if $\mathbf{y} \in \mathrm{Im}(\mathbf{H})$, then $P_{\mathrm{H}}\mathbf{y} = \mathbf{y}$.

SOLUTION If $\mathbf{y} \in \mathrm{Im}(\mathbf{H})$, there exists $\mathbf{x}$ such that $\mathbf{y} = \mathbf{Hx}$ and

$$P_{\mathrm{H}}\mathbf{y} = \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{Hx} = \mathbf{Hx} = \mathbf{y}.$$

$\square$

Problem 7: Projection with Mahalanobis distance

Show that the projection matrix $P_{\mathrm{H}} = \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T$ defined in class should be replaced by

$$P_{\mathrm{H}} = \mathbf{H}(\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W}$$

when dealing with weighted least square problem characterized by the symmetric positive definite matrix $\mathbf{W}$.

SOLUTION The normal equations in this case are $\mathbf{H}^T\mathbf{W}\mathbf{H}\mathbf{x} = \mathbf{H}^T\mathbf{W}\mathbf{y}$, from which we derive

$$\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W}\mathbf{y} \quad \text{and} \quad \hat{\mathbf{y}} = \mathbf{H}(\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W}\mathbf{y}.$$

We would like to think of $\mathbf{H}(\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W}$ as a projection matrix, but for what space?

It turns out that all we have to do is consider a new inner product defined as $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^T\mathbf{W}\mathbf{x}$. You can check that it satisfies all the properties of an inner product.

Note that $\hat{\mathbf{y}} \in \mathrm{Im}(\mathbf{H})$ by construction. In addition, for all $\mathbf{z} \in \mathrm{Im}(\mathbf{H})$, we have for some $\mathbf{x}$

$$(\mathbf{y} - \hat{\mathbf{y}})^T\mathbf{W}\mathbf{z} = \mathbf{y}^T(\mathbf{I} - \mathbf{W}\mathbf{H}(\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^T)\mathbf{W}\mathbf{H}\mathbf{x} = 0$$

so that the error $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to $\mathrm{Im}(\mathbf{H})$. Consequently, one can show that $\hat{\mathbf{y}}$ is indeed the orthogonal projection and $\mathbf{H}(\mathbf{H}^T\mathbf{W}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{W}$ is our project matrix. $\square$

Problem 8: Constrained least-square

We are now interested in solving the least-square problem $\min_x \|\mathbf{y} - \mathbf{Hx}\|_2^2$ subject to the linear constraint $\mathbf{Ax} = \mathbf{b}$. Note that this is different from a regularized least square. Using Lagrange multipliers, one can show that the constrained problem is equivalent to the minimization of the following unconstrained problem

$$\min_{x} \left(\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \boldsymbol{\lambda}^T (\mathbf{A}\mathbf{x} - \mathbf{b}) \right).$$

using the vector $\boldsymbol{\lambda}$ of Lagrange multipliers.

[Q1] Show that the stationarizing $\boldsymbol{\lambda}$ is

$$\boldsymbol{\lambda} = 2[\mathbf{A}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{A}^T]^{-1}(\mathbf{A}\mathbf{x}_{\mathrm{LS}} - \mathbf{b})$$

where $\mathbf{x}_{\mathrm{LS}}$ is the unconstrained least-square solution.

SOLUTION Taking the gradient with respect to $\boldsymbol{\lambda}$ and setting to zero we obtain the constraint $\mathbf{A}\mathbf{x} = \mathbf{b}$. Taking the gradient with respect to $\mathbf{x}$ and setting to zero, we obtain

$$\begin{aligned} & 2\mathbf{H}^T\mathbf{H}\mathbf{x} - 2\mathbf{H}^T\mathbf{y} + \mathbf{A}^T\boldsymbol{\lambda} = 0 \\ \Leftrightarrow & \mathbf{x} - \underbrace{(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{y}}_{x_{\mathrm{LS}}} + \frac{1}{2}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{A}^T\boldsymbol{\lambda} = 0 \\ \Leftrightarrow & \underbrace{\mathbf{A}\mathbf{x}}_{=b} - \mathbf{A}\mathbf{x}_{\mathrm{LS}} + \frac{1}{2}\mathbf{A}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{A}^T\boldsymbol{\lambda} = 0, \end{aligned}$$

i.e., $\boldsymbol{\lambda} = 2[\mathbf{A}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{A}^T]^{-1}(\mathbf{A}\mathbf{x}_{\mathrm{LS}} - \mathbf{b})$. □

[Q2] Conclude that the constrained least-square solution is

$$\mathbf{x}_c = \mathbf{x}_{\mathrm{LS}} - (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{A}^T[\mathbf{A}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{A}^T]^{-1}(\mathbf{A}\mathbf{x}_{\mathrm{LS}} - \mathbf{b})$$

SOLUTION Substituting the value of $\boldsymbol{\lambda}$ into our first equation we have

$$\begin{aligned} \mathbf{x} &= (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{y} - \frac{1}{2}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{A}^T\boldsymbol{\lambda} \\ &= \mathbf{x}_{\mathrm{LS}} - (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{A}^T[\mathbf{A}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{A}^T]^{-1}(\mathbf{A}\mathbf{x}_{\mathrm{LS}} - \mathbf{b}). \end{aligned}$$

□