CMDA 3606 · MATHEMATICAL MODELING II

Problem Set 4

Posted 13 February 2019. Due at 5pm on Thursday, 21 February 2018.

Basic guidelines: Students may discuss the problems on this assignment, but each student must submit his or her individual writeup and code. (In particular, you *must write up your own individual MATLAB code.*) Students may consult class notes and other online resources for general information; cite all your sources and list those with whom you have discussed the problems.

1. [15 points: 12 points for (a); 3 points for (b)]

   Recall the definition of the Euclidean norm (or 2-norm) of a vector $\|\mathbf{v}\|^2 = \mathbf{v}^T\mathbf{v}$ for any vector $\mathbf{v} \in \mathbb{R}^n$.

   In high school geometry, you might have learned this fact about parallelograms: *the sum of the squares of the lengths of the four sides equals the sum of the squares of the lengths of the two diagonals.*

   (a) Show that this holds in n-dimensional space by verifying the identity

   $$\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2.$$

   (b) Explain (with a simple picture in two-dimensional space) how the identity you proved in part (a) corresponds to the high-school statement above about parallelograms.

2. [15 points]

   In many situations, one has some subset $\mathcal{U}$ of $\mathbb{R}^n$, and wants a convenient way to talk about *the set of all vectors* $\mathbf{v} \in \mathbb{R}^n$ *that are orthogonal to* $\mathcal{U}$. This set is called the *orthogonal complement* of $\mathcal{U}$:

   $$\mathcal{U}^\perp = \{\mathbf{v} \in \mathbb{R}^n : \mathbf{u}^T\mathbf{v} = 0 \text{ for all } \mathbf{u} \in \mathcal{U}\}.$$

   Prove that $\mathcal{U}^\perp$ is a subspace of $\mathbb{R}^n$.

   (The *orthogonal complement* gives a very compact way to write the *Fundamental Theorem of Linear Algebra*: $\mathcal{R}(\mathbf{A})^\perp = \mathcal{N}(\mathbf{A}^T)$ and $\mathcal{N}(\mathbf{A})^\perp = \mathcal{R}(\mathbf{A}^T)$.)

3. [35 points: 5 points per part]

   Suppose you wish to construct a linear polynomial, $p(x) = \alpha + \beta x$ that approximates some function $f(x)$ at the points $x_1, x_2, x_3 \in \mathbb{R}$. This leads to the least squares problem $\min_{\mathbf{c}\in\mathbb{R}^2} \|\mathbf{b} - \mathbf{A}\mathbf{c}\|$, where

   $$\mathbf{A} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \end{bmatrix}.$$

   For this problem, suppose that $f(x) = \sqrt{1+x}$ and $x_1 = -1$, $x_2 = 0$, and $x_3 = 1$.

   (a) Compute (by hand) the pseudoinverse $\mathbf{A}^+ = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$.

   (b) Compute (by hand) the matrix $\mathbf{P} = \mathbf{A}\mathbf{A}^+$ and confirm (using the specific values in $\mathbf{P}$) that $\mathbf{P}^2 = \mathbf{P}$.

   (c) In MATLAB or by hand, compute the value of $\mathbf{c} = [\alpha,\ \beta]^T$ that solves the least squares problem.

   (d) Produce a plot in MATLAB showing how your linear approximation $p(x) = \alpha + \beta x$ compares to $f(x) = \sqrt{1+x}$, for x in the range $[-1, 2]$. The following commands give some helpful MATLAB syntax for plotting functions, which you could adapt to solve this problem.

```
g  = @(x) exp(x) + x.^2;                 % define some function of x
xx = linspace(-1,2,500);                 % define a fine grid of x points
plot(xx,g(xx),'b-','linewidth',2)        % plot g(x) as a thick blue line
hold on
plot(xx,1+xx+xx.^2,'r-','linewidth',2)   % plot 1+x+x^2 as a thick red line
```

(e) Suppose we now wish to approximate $f(x)$ with a quadratic polynomial $q(x) = \alpha+\beta x+\gamma x^2$, again at the points x_1, x_2, and x_3. Set this up as a least squares problem of the form $\min_{\mathbf{c}\in\mathbb{R}^3} \|\mathbf{Ac}-\mathbf{b}\|$, i.e., specify the appropriate entries in $\mathbf{A}$, $\mathbf{b}$, and $\mathbf{c}$. *Without solving the least squares problem*, what can you say about the minimal error $\displaystyle\min_{\mathbf{c}\in\mathbb{R}^3} \|\mathbf{Ac}-\mathbf{b}\|$? (Hint: think about the dimensions of $\mathbf{A}$, $\mathbf{c}$, and $\mathbf{b}$.)

(f) Suppose now that we wish to fit a cubic $r(x) = \alpha+\beta x+\gamma x^2+\delta x^3$ to the same data points x_1, x_2, and x_3. Set this up as a least squares problem, $\min_{\mathbf{c}\in\mathbb{R}^4} \|\mathbf{Ac}-\mathbf{b}\|$. Can you solve this problem to find α, β, γ, and δ? Is there a unique choice for these four parameters so that $r(x)$ optimally approximates $(x_1, f(x_1))$, $(x_2, f(x_2))$, and $(x_3, f(x_3))$?

(g) Produce a new MATLAB plot to show how your cubic approximation $r(x)$ from part (f) compares to $f(x) = \sqrt{1+x}$ for x in the range $[-1, 2]$. If you found a unique solution in (f), plot that single $r(x)$. If you found a nonunique solution, plot three distinct functions $r(x)$ that solve the least squares problem.

4. [35 points: 5 points for (a); 7 points each for (b) and (e); 8 points each for (c) and (d)]

Recall our earlier example of mechanical trusses at static equilibrium. Our four-step modeling process led to the equation $\mathbf{A}^T\mathbf{KAx} = \mathbf{f}$. Thus far we have considered situations where we knew how the structure was connected (encoded in $\mathbf{A}$), the stiffness of the struts (the diagonal of $\mathbf{K}$), and the loading force ($\mathbf{f}$), and we solved $(\mathbf{A}^T\mathbf{KA})\mathbf{x} = \mathbf{f}$ for the unknown displacements $\mathbf{x}$.

Many engineering applications present a different version of this problem: We can still see how the structure is connected, but we do not know the stiffness parameters for the struts: we do not know the values to put into $\mathbf{K}$. To find these parameters, we can apply different loads $\mathbf{f}$ to the structure, and *measure* the resulting displacements $\mathbf{x}$. Can such experiments reveal the stiffness values $k_1, \ldots, k_m$?

Consider this approach. Start with $\mathbf{A}^T\mathbf{KAx} = \mathbf{f}$. Let $\mathbf{c} = \mathbf{Ax}$, which we can compute if we are given $\mathbf{A}$ and then measure $\mathbf{x}$ for a given $\mathbf{f}$. This is the key observation:

$$\mathbf{KAx} = \mathbf{Kc} = \begin{bmatrix} k_1 & & \\ & \ddots & \\ & & k_m \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} k_1c_1 \\ \vdots \\ k_mc_m \end{bmatrix} = \begin{bmatrix} c_1 & & \\ & \ddots & \\ & & c_m \end{bmatrix} \begin{bmatrix} k_1 \\ \vdots \\ k_m \end{bmatrix} = \mathbf{Ck},$$

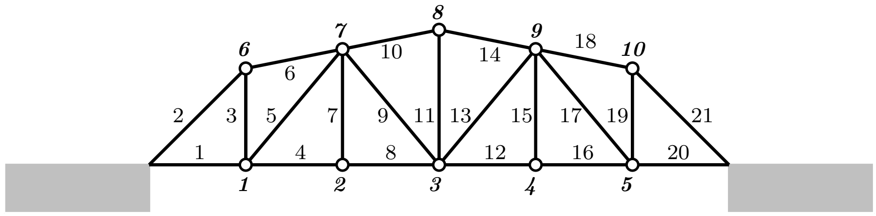
where $\mathbf{C} = \mathrm{diag}(\mathbf{c})$ is the diagonal matrix whose (j, j) entry equals c_j, and $\mathbf{k}$ is the vector with k_j in its jth position. Then $\mathbf{A}^T\mathbf{KAx} = \mathbf{f}$ is equivalent to

$$(\mathbf{A}^T\mathbf{C})\mathbf{k} = \mathbf{f},$$

which we hope to solve for the *unknown* stiffness values $\mathbf{k}$.

(a) If the truss has n nodes (hence $2n$ degrees of freedom for planar motion) and m struts, what size is the matrix $\mathbf{A}^T\mathbf{C}$?

(b) Suppose that $m > 2n$ (ensuring the truss is stable: after all, we are unlikely to perform this study on an unstable – collapsed – truss). What do your answer to part (a) and the Fundamental Theorem of Linear Algebra tell you about the solvability of $(\mathbf{A}^T\mathbf{C})\mathbf{k} = \mathbf{f}$ for the unknown $\mathbf{k}$? Specifically, given the size of $\mathbf{A}^T\mathbf{C}$, do you expect this equation to have a solution for all $\mathbf{f}$? Explain. (Assume that $\mathbf{x}$ is polluted with noise, so it has no special properties that ensure $\mathbf{A}^T\mathbf{Ck} = \mathbf{f}$ has a solution, beyond what can potentially be gleaned from the dimensions of $\mathbf{A}^T\mathbf{C}$.)



Suppose the bridge that appeared in Problem Set 3 supports train traffic crossing the New River. The struts are labeled with upright numbers (1, 2, ..., 21), the nodes with bold-italic numbers (***1***, ***2***, ..., ***10***), so, for example, f_{12} is the horizontal force applied to the top-left node, ***6***.

You have been hired to analyze whether this old bridge needs repair, by assessing the stiffness values $\mathbf{k}$ of these aging struts. You can experiment on the bridge by applying any desired load $\mathbf{f}$ to the structure, and measuring the corresponding vertical and horizontal displacements $\mathbf{x}$ of the nodes. Your instrumentation measures $\mathbf{x}$ to about a 1% relative error; this noise will compromise your results.

To simulate these measurements, you are given the `bridge.p` MATLAB code (on the class website). You collect data by creating a vector of loads `f` and calling `x = bridge(f)`. (So you are not tempted to peek at the code, we have compiled the MATLAB `.m` file as an executable but unreadable `.p` file.) The `bridge_A.mat` file contains the $\mathbf{A}$ matrix; download this file into your working directory and `load bridge_A` to get this `A` matrix in your MATLAB workspace.

(c) Construct a vector $\mathbf{f}$, compute $\mathbf{x}$ using `bridge.p`, construct $\mathbf{c} = \mathbf{A}\mathbf{x}$, and then solve the least squares problem

$$\min_{\mathbf{k}\in\mathbb{R}^m} \|(\mathbf{A}^T\mathbf{C})\mathbf{k} - \mathbf{f}\|$$

using MATLAB's \ command (`k = (A'*C)\f`). Explain the physical interpretation of your choice of $\mathbf{f}$ (Where did you choose to apply the forces?), and show the resulting estimated `k` values in a bar chart (MATLAB's `bar` command).

(d) The noise in the measurements can cause significant errors in the estimated value of $\mathbf{k}$. To obtain a better estimate, we can combine the results of several experiments. For example, if you conduct experiments with two different loads $\mathbf{f}^{(1)}$ and $\mathbf{f}^{(2)}$ giving measured displacements $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$, then you hope

$$\mathbf{A}^T\mathbf{C}^{(1)}\mathbf{k} = \mathbf{f}^{(1)}, \qquad \mathbf{A}^T\mathbf{C}^{(2)}\mathbf{k} = \mathbf{f}^{(2)},$$

where $\mathbf{C}^{(1)} = \mathrm{diag}(\mathbf{A}\mathbf{x}^{(1)})$ and $\mathbf{C}^{(2)} = \mathrm{diag}(\mathbf{A}\mathbf{x}^{(2)})$. You can find the $\mathbf{k}$ that best satisfies these two equations simultaneously as

$$\min_{\mathbf{k}\in\mathbb{R}^m} \left\| \begin{bmatrix} \mathbf{A}^T\mathbf{C}^{(1)} \\ \mathbf{A}^T\mathbf{C}^{(2)} \end{bmatrix} \mathbf{k} - \begin{bmatrix} \mathbf{f}^{(1)} \\ \mathbf{f}^{(2)} \end{bmatrix} \right\|.$$

Following this example, combine the measurements $\mathbf{x}^{(j)}$ obtained from `bridge.p` for *at least three* different loads $\mathbf{f}^{(j)}$ to determine an optimal $\mathbf{k}$. Again describe the physical meaning of your choice of $\mathbf{f}^{(j)}$, and show your estimated $\mathbf{k}$ using a `bar` chart. Do your results look more reasonable than those obtained in part (c)?

(e) Suppose that you should consider the bridge unsafe if any of the struts have a stiffness value k_j smaller than 0.25. Is the bridge faulty? If so, name the problematic strut(s). The bridge will be very expensive to repair, so if you find it unsafe, explain why you have confidence that your findings are accurate, and are not badly affected by the error in your measurements.

This set-up describes the first "inverse problem" we have seen this semester: using experiments and measurements to infer the material properties of a system. It is adapted from a tissue biopsy example of Steve Cox.

5. [6 bonus points: 3 for part (a), 3 for part (b)].

Suppose $\mathbf{P} \in \mathbb{R}^{n\times n}$ is a *projector*, meaning that $\mathbf{P}^2 = \mathbf{P}$.

(a) Show that if $\mathbf{P} = \mathbf{P}^T$, then $\|\mathbf{P}\mathbf{x}\| \le \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^n$.

(b) Show that if $\mathbf{P} \ne \mathbf{P}^T$, then there exists some $\mathbf{x} \in \mathbb{R}^n$ for which $\|\mathbf{P}\mathbf{x}\| > \|\mathbf{x}\|$.