

MATH 437 Homework 6 (20 points)

1. (2 points) Find the linear least-squares polynomial approximation on $[-1, 1]$ to

$$f(x) = x^3.$$

Solution. Let $p(a, b; x) = a + bx$, where a and b will be determined so that p is the linear least-squares approximation to f . The least-squares approximation minimizes the error

$$e(a, b) := \int_{-1}^1 (f(x) - p(a, b; x))^2 dx.$$

Thus, the partial derivatives of e with respect to a and b should be 0:

$$\begin{aligned}\partial_a e &= -2 \int_{-1}^1 (f(x) - p(a, b; x)) \partial_a p(a, b; x) dx \\ &= -2 \int_{-1}^1 x^3 - (a + bx) dx \\ &= 2a = 0, \\ \partial_b e &= -2 \int_{-1}^1 (f(x) - p(a, b; x)) \partial_b p(a, b; x) dx \\ &= -2 \int_{-1}^1 (x^3 - (a + bx))x dx \\ &= -2(2/5 - 2b/3) = 0.\end{aligned}$$

Therefore, $a = 0$ and $b = 3/5$. □

Notice that requiring the partial derivatives vanish is equivalent to requiring the difference

$$g(a, b; x) := f(x) - p(a, b; x)$$

is orthogonal to the linear basis monomials 1 and x :

$$\begin{aligned}\partial_a e = 0 &\iff \int_{-1}^1 g(a, b; x) dx = 0, \\ \partial_b e = 0 &\iff \int_{-1}^1 g(a, b; x)x dx = 0.\end{aligned}$$

This generalizes. To find the degree n least-squares approximation,

$$p(a_0, \dots, a_n; x) := \sum_{i=0}^n a_i x^i$$

to a function $f(x)$ on an interval $[x_0, x_1]$, the difference

$$g(a_0, \dots, a_n; x) := f(x) - p(a_0, \dots, a_n; x)$$

must be orthogonal to the basis monomials $1, x, \dots, x^n$:

$$\int_{x_0}^{x_1} g(a_0, \dots, a_n; x) x^i dx = 0 \quad \text{for all } i = 0, \dots, n.$$

This gives an $n \times n$ linear system for the coefficients a_i .

2. (3 points) Use the Gram–Schmidt process to construct an orthogonal basis $\phi_0(x), \phi_1(x), \phi_2(x)$ for the space of polynomials of degree 2 on the interval $[0, 2]$.

Solution. For any functions f, g on $[0, 2]$, we let

$$(f, g) := \int_0^2 f(x)g(x) dx.$$

We start with the non-orthogonal monomial basis $\psi_0(x) = 1, \psi_1(x) = x, \psi_2(x) = x^2$, and we orthogonalize via Gram–Schmidt. We set $\phi_0 := \psi_0 = 1$. Then, we project ψ_1 onto $\text{span}\{\phi_0\}$:

$$p_1 := \frac{(\psi_1, \phi_0)}{(\phi_0, \phi_0)} \phi_0 = \frac{\int_0^2 x dx}{\int_0^2 1 dx} 1 = 1.$$

By removing p_1 from ψ_1 , we obtain ϕ_1 :

$$\phi_1 := \psi_1 - p_1 = x - 1.$$

We verify that ϕ_1 and ϕ_0 are orthogonal:

$$(\phi_1, \phi_0) = \int_0^2 x - 1 dx = 0.$$

To construct ϕ_2 , we project ψ_2 onto $\text{span}\{\phi_0, \phi_1\}$:

$$\begin{aligned} p_2 &:= \frac{(\psi_2, \phi_0)}{(\phi_0, \phi_0)} \phi_0 + \frac{(\psi_2, \phi_1)}{(\phi_1, \phi_1)} \phi_1 \\ &= \frac{\int_0^2 x^2 dx}{\int_0^2 1 dx} 1 + \frac{\int_0^2 x^2(x-1) dx}{\int_0^2 (x-1)^2 dx} (x-1) \\ &= 4/3 + 2(x-1) = 2x - 2/3. \end{aligned}$$

By removing this from ψ_2 , we obtain ϕ_2 :

$$\phi_2 := \psi_2 - p_2 = x^2 - 2x + 2/3.$$

We verify that ϕ_2 is orthogonal to ϕ_0 and ϕ_1 :

$$\begin{aligned} (\phi_2, \phi_0) &= \int_0^2 x^2 - 2x + 2/3 dx = 0, \\ (\phi_2, \phi_1) &= \int_0^2 (x^2 - 2x + 2/3)(x-1) dx = 0. \end{aligned}$$

Thus, $\phi_0 = 1, \phi_1 = x - 1, \phi_2 = x^2 - 2x + 2/3$ is an orthogonal basis. □

3. (3 points) Compute the first three iterations of the power method without normalization (see notes) with the following matrix:

$$A := \begin{bmatrix} 4 & 2 & 1 \\ 2 & 4 & 0 \\ 1 & 0 & 4 \end{bmatrix}.$$

Start with initial eigenvector $\mathbf{x}_0 = (1, 1, 1)^T$. Report the approximate eigenvalue and eigenvector (μ_i, \mathbf{x}_i) for $i = 1, 2, 3$.

Solution. We follow the standard power method without normalization:

$$x_{k+1} = Ax_k, \quad \mu_{k+1} := \frac{x_{k+1,p}}{x_{k,p}},$$

where p is such that all $x_{k,p}$ are (eventually) nonzero. See `problem_3.py` in the code solutions. The exact sequence of eigenvalues depends on if you use the first, second, or third coordinate. The sequence of eigenvectors will be the same in all cases.

With normalization, the sequence of eigenvectors is divided by its largest component at each step. I accepted this answer as well.

```

first coordinate
7.0 [7 6 5]
6.428571428571429 [45 38 27]
6.288888888888889 [283 242 153]

second coordinate
6.0 [7 6 5]
6.333333333333333 [45 38 27]
6.368421052631579 [283 242 153]

third coordinate
5.0 [7 6 5]
5.4 [45 38 27]
5.666666666666667 [283 242 153]
```

□

4. (3 points) Let A be an $n \times n$ matrix with n distinct eigenvalues ordered such that

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$$

and with corresponding eigenvectors v_1, \dots, v_n . Let $\beta_2, \dots, \beta_n \in \mathbb{R}$, where $\beta_2 \neq 0$, and set

$$x_0 := \sum_{i=2}^n \beta_i v_i,$$

so that x_0 is in the span of all but the first eigenvectors. Prove that the sequence of approximate eigenvalues μ_k obtained from the power method starting from this initial condition converges to the second-largest eigenvalue λ_2 .

Solution. The power method (without normalization) to construct the approximating sequence of eigenvectors x_k is

$$x_k := Ax_{k-1} = \cdots = A^k x_0 = \sum_{i=2}^n \beta_i \lambda_i^k v_i = \beta_2 \lambda_2^k \left(v_2 + \sum_{i=3}^n \frac{\beta_i}{\beta_2} \left(\frac{\lambda_i}{\lambda_2} \right)^k v_i \right),$$

assuming $\beta_2 \neq 0$. Let

$$r_k := v_2 + \sum_{i=3}^n \frac{\beta_i}{\beta_2} \left(\frac{\lambda_i}{\lambda_2} \right)^k v_i.$$

Since $|\lambda_i|/|\lambda_2| < 1$ when $i > 2$, we observe that $r_k \rightarrow v_2$ as $k \rightarrow \infty$. Thus,

$$\frac{x_k}{\lambda_2^k} = \beta_2 r_k \rightarrow \beta_2 v_2$$

as $k \rightarrow \infty$. Since v_2 is an eigenvector, some component $v_{2,p} \neq 0$. Thus, for k large enough, all $r_{k,p} \neq 0$. Since $\beta_2 \neq 0$ and $\lambda_2 \neq 0$, it also follows that eventually $x_{k,p} \neq 0$. Therefore, by taking the p th component and dividing:

$$\frac{x_{k+1,p}}{x_{k,p}} = \lambda_2 \frac{r_{k+1,p}}{r_{k,p}} \rightarrow \lambda_2$$

as $k \rightarrow \infty$. The approximate eigenvalues for the power iteration are defined as

$$\mu_{k+1} := \frac{x_{k+1,p}}{x_{k,p}},$$

and we just proved that they converge to λ_2 . □

5. (3 points) Use the Gershgorin Circle Theorem to find bounds on the eigenvalues for

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}.$$

Solution. See Theorem 9.1 in the textbook. The eigenvalues lie within the union of the *Gershgorin circles* in the complex plane:

$$R_i := \{z \in \mathbb{C} \mid |z - A_{i,i}| \leq \sum_{j \neq i} |A_{i,j}|\}.$$

For the given matrix:

$$R_1 = \{z \in \mathbb{C} \mid |z - 2| \leq 1\},$$

$$R_2 = \{z \in \mathbb{C} \mid |z - 2| \leq 0\},$$

$$R_3 = \{z \in \mathbb{C} \mid |z - 2| \leq 1\}.$$

Thus, all eigenvalues λ satisfy $|\lambda - 2| \leq 1$. Since the matrix is real and symmetric, the eigenvalues are actually all real-valued, so we conclude they lie in the interval $[1, 3]$. □

6. For the following set of vectors:

(a) (1 point) Show that the set is linearly independent.

(b) (2 points) Use the Gram-Schmidt process to orthogonalize the set.

$$v_1 = (1, 1, 1, 1)^T, \quad v_2 = (0, 2, 2, 2)^T, \quad v_3 = (1, 0, 0, 1)^T.$$

Solution. (a) Suppose $c_1 v_1 + c_2 v_2 + c_3 v_3 = 0$. Then,

$$c_1 + c_3 = 0,$$

$$c_1 + 2c_2 = 0,$$

$$c_1 + 2c_2 + c_3 = 0.$$

Using the second equation in the last, $c_3 = 0$. Then, from the first, $c_1 = 0$. Then, from the second, $c_2 = 0$.

(b) We let $(v, w) := \sum_i v_i w_i$ denote the standard dot product of vectors. We let ϕ_1, ϕ_2, ϕ_3 denote the desired orthogonal vectors. We set $\phi_1 = v_1$. Then, we project v_2 onto $\text{span } \phi_1$:

$$p_2 := \frac{(v_2, \phi_1)}{(\phi_1, \phi_1)} \phi_1 = \frac{3}{2} \phi_1.$$

Then, we remove this from v_2 to get ϕ_2 :

$$\phi_2 = v_2 - p_2 = (-3/2, 1/2, 1/2, 1/2)^T.$$

Now, we project v_3 onto $\text{span}\{\phi_1, \phi_2\}$:

$$\begin{aligned} p_3 &= \frac{(v_3, \phi_1)}{(\phi_1, \phi_1)}\phi_1 + \frac{(v_3, \phi_2)}{(\phi_2, \phi_2)}\phi_2, \\ &= \frac{1}{2}\phi_1 - \frac{1}{3}\phi_2. \end{aligned}$$

Then, we subtract from v_3 to obtain ϕ_3 :

$$\phi_3 = v_3 - p_3 = (0, -1/3, -1/3, 2/3)^T.$$

□

7. (3 points) Construct an orthogonal matrix Q such that $Q^T A Q = D$, where D is a diagonal matrix and

$$A = \begin{bmatrix} 4 & 2 & 1 \\ 2 & 4 & 0 \\ 1 & 0 & 4 \end{bmatrix}$$

Solution. Since A is symmetric, the orthogonal matrix Q that we seek has columns consisting of normalized eigenvectors of A , and the diagonal matrix has entries consisting of the eigenvalues of A (see Corollary 9.17 in the textbook). We first find the eigenvalues:

$$\det(A - \lambda I) = (4 - \lambda)^3 - 4(4 - \lambda) - (4 - \lambda) = 0.$$

Thus, $\lambda = 4$ or

$$(4 - \lambda)^2 - 5 = 0,$$

so $\lambda = 4 \pm \sqrt{5}$. Thus, the diagonal matrix D is

$$D = \begin{bmatrix} 4 - \sqrt{5} & & \\ & 4 & \\ & & 4 + \sqrt{5} \end{bmatrix}.$$

Now, we find corresponding eigenvectors and then normalize. For a general eigenvector $x = (x_1, x_2, x_3)^T$,

$$Ax = (4x_1 + 2x_2 + x_3, 2x_1 + 4x_2, x_1 + 4x_3)^T = \lambda x.$$

Thus,

$$\begin{aligned} 4x_1 + 2x_2 + x_3 &= \lambda x_1, \\ 2x_1 + 4x_2 &= \lambda x_2, \\ x_1 + 4x_3 &= \lambda x_3. \end{aligned}$$

For $\lambda = 4$, solving this system gives

$$x = x_2(0, 1, -2)^T$$

as an eigenvector, where x_2 is arbitrary. By normalizing so that $\|x\|_2 = 1$, we conclude that $x_2 = 1/\sqrt{5}$.

For $\lambda = 4 - \sqrt{5}$, we get

$$x = x_1(\sqrt{5}, -2, -1)^T$$

as a family of eigenvectors. To normalize, we set $x_1 = 1/\sqrt{10}$. For $\lambda = 4 + \sqrt{5}$, we get

$$x = x_1(\sqrt{5}, 2, 1)^T$$

as a family of eigenvectors. To normalize, we set $x_1 = 1/\sqrt{10}$. Thus, the matrix Q is formed by setting the columns to the normalized eigenvectors:

$$Q = \begin{bmatrix} \sqrt{5}/\sqrt{10} & 0 & \sqrt{5}/\sqrt{10} \\ -2/\sqrt{10} & 1/\sqrt{5} & 2/\sqrt{10} \\ -1/\sqrt{10} & -2/\sqrt{5} & 1/\sqrt{10} \end{bmatrix}.$$

□