

Assignment 4, due October 14, before class

Please follow [instructions on the assignments page](#)

About this assignment: Part 1: intuition about eigenvalues and singular values, linear algebra is arguably the most important part of applied math. Part 2: computer algorithms don't always do what you expect if you don't remember that computer arithmetic is not done exactly.

$$y = \arg \min_x f(x)$$

1. (*Geometry of quadratic curves in 2D*) A general quadratic form in 2D has the form

$$Q(x, y) = ax^2 + 2bxy + cy^2 .$$

For this problem, we write the coordinates of a point in the plane as $\begin{pmatrix} x \\ y \end{pmatrix}$ when we do linear algebra and (x, y) when we do geometry.

- (a) Find the symmetric 2×2 matrix A so that

$$(x, y) A \begin{pmatrix} x \\ y \end{pmatrix} = Q(x, y) .$$

The entries of A involve the numbers a, b, c . The 2 of $2bxy$ in the definition of Q simplifies the form of A . The quadratic form is *degenerate* if the corresponding matrix A is singular. Otherwise it is *non-degenerate*.

- (b) We define the *level set* and the *circle image* as¹

$$L_\rho = \{ (x, y) \mid Q(x, y) = \rho \} \quad , \quad C_\rho = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} = A \begin{pmatrix} \xi \\ \eta \end{pmatrix} , \quad \xi^2 + \eta^2 = \rho \right\} .$$

You can see (not to hand in) that each ρ level sets is “similar” either to the $\rho = 1$ level set or the $\rho = -1$ level set (if $\rho \neq 0$), and the ρ circle images are all similar to the $\rho = 1$ circle image (if $\rho \neq 0$). Therefore, we will always take $\rho = 1$ and write L and C for L_1 and C_1 . The set L_ρ might be empty ($Q(x, y) = \rho$ is impossible). To rule that out, we take $\rho = 1$ (as above) and $a > 0$. Show that $(0, 0) \notin L$ if Q is non-degenerate. Use the method of Lagrange multipliers to find a characterization of $u = (u_1, u_2)$, which is the closest point on L to the origin

$$\min_{(x, y) \in L} x^2 + y^2 \quad , \quad u_1 = (v_1, w_2) = \arg \min_{(x, y) \in L} x^2 + y^2 .$$

Show that the optimality condition implies that u_1 is an eigenvector of A .

- (c) The quadratic form is called *elliptic* if L is bounded. In that case there is a well defined maximum

$$\max_{(x, y) \in L} x^2 + y^2 .$$

¹Math people think of ξ and η as Greek versions of x and y , but I don't imagine people from Greece think that. A level set of a function is the set of points where the function takes a given value. The term comes from contour maps, where the function is the height. The image of a set under a function is the set of values the function takes when its argument is in the given set.

Show that if Q is elliptic, the maximizer u_2 is another eigenvector. Show that both eigenvalues of A are positive and Q is positive definite in this case. The level set for a positive definite quadratic form is called an *ellipse*. The minimizer direction u_1 is the *minor axis* and the maximizer direction u_2 is the *major axis*.

- (d) For the case $Q = 2x^2 - 2xy + y^2$, find the maximizer and minimizer explicitly and verify by direct calculation that u_1 is perpendicular to u_2 .
 - (e) If Q is non-degenerate but not positive definite, we say Q is *indefinite* or *hyperbolic*. Let u_2 be a unit vector perpendicular to u_1 . Show that if u_1 is a minimizer, then Q is *diagonal* in the u_1, u_2 basis, in the sense that if $(x, y) = \xi u_1 + \eta u_2$, then $Q(x, y) = \alpha_1 \xi^2 + \alpha_2 \eta^2$. (This makes the corresponding matrix a diagonal matrix.). How are α_1 and α_2 related to the eigenvalues of A , λ_1 and λ_2 ? Show that the u_2 direction is an *asymptote* of L in the sense that L goes to infinity while getting closer and closer to the u_2 direction. The curve L is called a *hyperbola* if Q is hyperbolic.
 - (f) Show that the circle image C is an ellipse whenever A is non-degenerate and that u_1, u_2 are the major and minor axes of this ellipse (in some order). You can do the “ellipse” part with or without algebra. The image of any bounded set (such as a circle) under a continuous function (such as a linear transformation) is a bounded set. That’s why it can’t be a hyperbola.
 - (g) Show that the circle image is an ellipse even if A is not symmetric.
2. (*Variational SVD*) The variational approach to the SVD is an adaptation of the variational approach to the symmetric eigenvalue problem. An $n \times m$ matrix, A , determines a mapping from m component vectors to n component vectors. This is expressed in various ways, depending on the point you’re trying to make:

$$y = Ax, \quad \begin{array}{ccc} \mathbb{R}^m & \xrightarrow{A} & \mathbb{R}^n \\ x & \rightsquigarrow & Ax \end{array} .$$

The image of the unit sphere under A is like the circle set from Exercise 1:

$$C = \left\{ Ax \mid \|x\|_2^2 = 1 \right\} .$$

Define vectors $v_1 \in \mathbb{R}^m$, $u_1 \in \mathbb{R}^n$, and the positive scale factor σ_1 by

$$\begin{aligned} \sigma_1 u_1 &= \arg \max_{x \in C} \|x\|_2, \quad \|u_1\|_2 = 1 \\ v_1 &= \arg \max_{y \in \mathbb{R}^m, \|y\|=1} \|Ay\|_2 \\ \sigma_1 &= \|Av_1\|_2 . \end{aligned}$$

Many of these formulas could have used $\|\cdot\|_2^2$ instead of $\|\cdot\|_2$ because maximizing a positive function is the same as maximizing the square, and requiring a positive number be equal to one is the same as requiring its square be equal to one. This defines σ_1 as the maximum stretch factor of the linear transformation A . We saw in class that you find σ_2 by maximizing with the orthogonality constraint $y^T v_1 = 0$. An orthogonality principle (from class) implies that the constrained result u_2 is also orthogonal to u_1 . This argument is asymmetric, treating orthogonality in the “source space”, $\mathbb{R}^m$ and the “target space”, $\mathbb{R}^n$ differently.

- (a) Suppose $x \in \mathbb{R}^n$ is any vector. Show that

$$\|x\|_2 = \max_{z \in \mathbb{R}^n, \|z\|_2=1} z^T x .$$

You may assume the *Cauchy Schwarz inequality*, for any two vectors in $\mathbb{R}^n$,

$$z^T x \leq \|z\|_2 \|x\|_2 .$$

This is trivial if $z^T x < 0$, because $\|z\|_2 \|x\|_2 \geq 0$. In this case $(-z)^T x = |z^T x| > 0$ and $\| -z \|_2 = \|z\|_2$. This proves the inequality with the absolute value on the right, as you might find it in Wikipedia (or from Claude, who got it from Wikipedia): $|z^T x| \leq \|z\|_2 \|x\|_2$. *Hint.* The solution has two parts. You must show that $\|x\|_2 \leq \|z\|_2 \|x\|_2$ for any allowed z , and that there is an allowed z with $\|x\|_2 = z^T x$.

(b) Show that

$$\begin{aligned} \sigma_1 &= \max_{\|z\|_1=1, \|w\|_2=1} z^T A w \\ &= \max_{\|z\|_1=1, \|w\|_2=1} w^T A^T z . \end{aligned}$$

Hint. This has two parts. First show the first formula, then show the right sides of the two lines are the same.

(c) Show that if we maximize $\|Ay\|_2$ with $\|y\|_2 = 1$ and $y^T v_1 = 0$, we get the same σ_2 that we would get by maximizing $\|A^T x\|_2$ with $\|x\|_2 = 1$ and $x^T u_1 = 0$.

3. (*Deflation*) Suppose we have a symmetric $n \times n$ matrix and an eigenvector $Au = \lambda u$. *Deflation* is the process of finding an $(n-1) \times (n-1)$ matrix whose eigenvalues are the rest of the eigenvalues of A . This is done by finding a basis in which A has the form

$$\tilde{A} = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & * & \cdots & * \\ \vdots & * & & * \\ 0 & * & \cdots & * \end{pmatrix} . \quad (1)$$

It is common when describing linear algebra algorithms to write $*$ for a number that does not have to be zero. This puts A into the block form

$$\tilde{A} = \begin{pmatrix} \lambda & 0 \\ 0 & B \end{pmatrix} .$$

The $n-1$ eigenvalues of the symmetric matrix B , together with λ , are the n eigenvalues of A .

- (a) Let $y \in \mathbb{R}^n$ be a unit norm vector ($\|y\|_2^2 = y^T y = 1$). The corresponding *Householder matrix* is

$$H = I - 2yy^T .$$

Show that H is symmetric and orthogonal.

- (b) This H is sometimes called the Householder *reflection*. Justify this by showing that $H^2 = I$ (reflect and reflect back and you're where you started) and that H represents reflection normal to the hyperplane normal to y in the sense that $y \xrightarrow{H} -y$ and $H z = z$ if z is in the normal hyperplane.

- (c) Show that if u is a unit norm vector, then there is a y so that u is the first column of H . *Hint.* Choose the entries of y one by one, starting with $y_1 = \frac{1}{2}\sqrt{1-u_1}$. Figure out why $\sqrt{1-u_1}$ is real.

- (d) Show that the matrix

$$\tilde{A} = H^T A H$$

Has the block diagonal form (1). *Hint.* H is symmetric, so the transpose is irrelevant. Let w_j be column j of H . The entries of $\tilde{A}$ are

$$\tilde{a}_{jk} = w_j^T A w_k .$$

Since $w_1 = u$ (by construction), the eigenvector orthogonality implies that the entries in the first row and column are zero, except $\tilde{a}_{11} = \lambda$.

(e) Show that if $Bv = \mu v$ (in $\mathbb{R}^{n-1}$) then

$$u_2 = \begin{pmatrix} 0 \\ v \end{pmatrix}$$

is an eigenvector of $\tilde{A}$ with eigenvalue μ . Find a corresponding eigenvector of A with eigenvalue μ .

(f) Show that if you have an algorithm that finds an eigenvector of A , it can be used with deflation to find all the eigenvectors and eigenvalues of A .

4. (*Power method*) “The” *power method* refers to several algorithms for finding an eigenvalue and eigenvector of a symmetric matrix. It uses the iteration

$$x_{k+1} = \frac{1}{Z_k} A x_k . \quad (2)$$

The normalization factors Z_k are chosen so that $\|x_{k+1}\|_2 = 1$. You first compute the matrix vector product $\tilde{x}_{k+1} = A x_k$, then find $Z_k = \|\tilde{x}_{k+1}\|_2$. Suppose the eigenvectors and eigenvalues satisfy $A u_j = \lambda_j u_j$ with $0 < \lambda_1 < \lambda_2 < \dots < \lambda_n$ (the eigenvalues are in increasing order and assumed to be distinct, and A is positive definite). The iteration (2) starts with an *initial guess* or *starting vector* x_0 . Show that if the u_n component of x_0 is not zero, then x_k converges to a multiple of u_n as $n \rightarrow \infty$. This means that if

$$x_0 = \sum_{j=1}^n a_{0,j} u_j .$$

and if $a_{0,n} \neq 0$, and if $\|u_n\|_2 = 1$, then $a_{j,n} \rightarrow \pm 1$ and $a_{k,j} \rightarrow 0$ for $j < n$, as $k \rightarrow \infty$. *Comments.*

- The “rate of convergence” (how fast it converges once it’s at all close) is determined by the *spectral gap*, $\rho = \frac{\lambda_n - \lambda_{n-1}}{\lambda_n}$. The *spectrum* of a matrix is the set of eigenvalues. The spectral gap is a gap in the spectrum, normalized to be dimensionless.
- How long it takes for x_k to be anything like u_n depends on the “overlap” between the eigenvector and the initial gap, which is $a_{0,n}$.
- If A is not positive definite, then x_k converges either to x_1 or x_n depending on whether or not $|\lambda_1| > |\lambda_n|$ or not.
- In the computer, you typically get $x_k \rightarrow \lambda_n$ even if $a_{0,n} = 0$ because of *roundoff error* (the fact that computer arithmetic is not exact). See exercise 5.
- This “vanilla” power method is not the best way to find eigenvalue and eigenvectors of a symmetric matrix. The codes in the Python libraries `numpy.linalg` and `scipy` are much better. It’s almost always better to use a package than to code something yourself, if it’s numerical linear algebra.

5. (*Loss of orthogonality*) Let A be the $d \times d$ symmetric *tridiagonal* matrix with 2 on the diagonal and -1 on the nearest subdiagonal and super-diagonal. *Tridiagonal* means that the only non-zeros are on the diagonal and on the immediate subdiagonal and superdiagonal. This means $a_{jj} = 2$, $a_{j,j-1} = a_{j,j+1} = -1$ and $a_{jk} = 0$ if $|k - j| > 1$.

(a) Find the variational definition of A by showing that $x^T A x = Q(x)$, where

$$Q(x_1, \dots, x_d) = \sum_{k=1}^{d+1} (x_k - x_{k-1})^2 . \quad (3)$$

This assumes *boundary conditions* $x_0 = 0$ and $x_{d+1} = 0$. The formula (3) has a simple form, but also can be written as

$$Q(x_1, \dots, x_d) = x_1^2 + \sum_{k=2}^d (x_k - x_{k-1})^2 + x_d^2 .$$

Use this to show that A is positive definite. *Warning.* x_k here is component k of x , not iterate k . x_n will mean iterate n in every other part of this exercise.

- (b) Write a code to do the power method on this matrix. Experiment with different n (starting from small to quite large) and different choices of initial guess (x_0). You can get an idea how close x_n is to being an eigenvector by estimating the eigenvalue using the Rayleigh quotient (writing μ_n for the estimate of λ_d found from x_n)

$$\mu_n = x_n^T A x_n .$$

If x_n is an eigenvector then μ_n is the corresponding eigenvalue, assuming x_n is normalized. Then you can calculate the residual $r_n = \mu_n x_n - A x_n$ and see how fast that goes to zero. Make a plot of $\|r_n\|_2$ as a function of n . It is probably more helpful to make a *semilog* plot, so that exponential convergence corresponds to a straight line. The convergence should become “more exponential” as n grown but is never exactly exponential. If you feel ambitious (but not to hand in unless you find a way to embed a movie in a PDF), make a movie of x_n as a function of n (your AI knows how to make the movie, but nobody knows how to embed it into a PDF). Try to see which x_0 lead to fast or slow convergence, which will be hard to do unless d is large.

- (c) Show (with a written argument, not computational demonstration) that if x_0 is orthogonal to u_d , then the iterates x_n are also orthogonal to u_d and “probably” converge to u_{d-1} (the eigenvector with second largest eigenvalue).
- (d) Once part (b) gives you u_d with $A u_d = \lambda_d u_d$ (not exactly, but to *machine precision*), find a vector x_0 orthogonal to this and do the power method. Does x_n converge to u_{d-1} as suggested by part (c) or to u_d ? Can you explain this behavior, which is called *loss of orthogonality*, as x_n stops being (approximately) orthogonal to u_d as n increases. Why? It may be interesting to plot μ_n as a function of n for large d . The algorithms might (if conditions are right) “try to converge” to λ_{d-1} before losing orthogonality and going to λ_d instead.
- (e) As an alternative to the method of part (d), which is one-time orthogonalization, compute the deflated matrix described in Exercise 3 and apply the power method to that. Do you get λ_d or λ_{d-1} ?
- (f) You *orthogonalize* a vector y against a vector u by subtracting the u component of y :

$$z = y - c u , \quad u^T z = 0 .$$

Find the formula for c , which is different depending on whether u is normalized or not.

- (g) The power method *with re-orthogonalization* for u_{d-1} does a re-orthogonalization every step

$$y_{n+1} = Ay_n \quad (\text{power method step})$$

$$z_{n+1} = y_{n+1} - c_{n+1}u_d \quad (\text{reorthogonalization})$$

$$x_{n+1} = \frac{1}{Z_{n+1}} z_{n+1} \quad (\text{renormalization}) .$$

Try it.