MATH-UA 253/MA-UY 3204 - Fall 2022 - Midterm (take-home)

Out of 20 points.

Problem 1 (analysis of gradient descent applied to a quadratic form): 6 points. Let $A \in \mathbb{R}^{n \times n}$ be a positive definite matrix, let $b \in \mathbb{R}^n$, and let $c \in \mathbb{R}$. Consider the quadratic form:

$$f(x) = \frac{1}{2} x^\top A x - b^\top x + c. \tag{1}$$

Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$ be the eigenvalues of A, and let v_i $(i = 1, \ldots, n)$ be the eigenvectors associated to each eigenvalue, normalized so that $\{v_i\}_{i=1}^n$ forms an orthonormal basis for $\mathbb{R}^n$.

In this problem, we'll analyze the case of minimizing f using gradient descent. That is, we seek to minimize f with the iteration:

$$x_{k+1} = x_k + \alpha p_k, \qquad p_k = -\nabla f(x_k), \tag{2}$$

where $\alpha > 0$.

1. Let x^* be the minimizer of f. Show that $x^* = A^{-1}b$, and that x^* is in fact the unique global minimum of f.
2. Show that $x_{k+1} - x^* = (I - \alpha A)(x_k - x^*)$ for all $k \geq 0$.
3. Now, assume that $\alpha = \lambda_1^{-1}$. Deduce that:

$$\|x_k - x^*\| \leq \left(1 - \frac{\lambda_n}{\lambda_1}\right)^k \|x_0 - x^*\|, \qquad k \geq 0. \tag{3}$$

4. Let $\alpha_i^{(k)}$ be the coefficient of $x_k - x^*$ in the $\{v_i\}_{i=1}^n$ basis corresponding to the ith basis vector, i.e.:

$$\alpha_i^{(k)} = v_i^\top (x_k - x^*). \tag{4}$$

   Express $\alpha_i^{(k)}$ in terms of k, λ_i, λ_1, and $\alpha_i^{(0)}$.

5. Using the solution to the previous problem, justify the following statement: *Gradient descent converges towards the minimizer faster in directions given by the eigenvectors of the Hessian of f corresponding to large eigenvalues than in directions with smaller eigenvalues.*
6. Finally, show that the distance to optimality at the kth step is given exactly by:

$$\|x_k - x^*\|_2^2 = \sum_{i=1}^n \left(1 - \frac{\lambda_i}{\lambda_1}\right)^{2k} \left(v_k^\top (x_0 - x^*)\right)^2. \tag{5}$$

Problem 2 (electric circuits): 5 points. Consider an electric circuit with $n+1$ nodes, $x_0, \dots, x_n$. If there is a resistor connecting nodes x_i and x_j, its resistance is denoted by $r_{ij} > 0$. The conductance between nodes x_i and x_j is then $\sigma_{ij} = 1/r_{ij}$. If x_i and x_j are not connected, $r_{ij} = \infty$. Assume that there is a path connecting nodes x_0 and x_n, and that we connect nodes x_0 and x_n to a battery to create a voltage difference of V across them. How can we find the voltage difference across the other nodes, and the currents across the resistors?

Let $v_0, \dots, v_n$ be the voltage of each node. The solution of this problem can be cast as an optimization problem to minimize the electrostatic energy of the system:

$$\text{minimize} \qquad e(v_0, \dots, v_n) = \frac{1}{2} \sum_{0 \le i < j \le n} \sigma_{ij} \cdot (v_i - v_j)^2. \tag{6}$$

1. Explain how to compute v_i for each node.

2. Ohm's law says that the current across each resistor is given by $\iota_{ij} = (v_i - v_j)/r_{ij}$. Show that *Kirchoff's law* holds:
$$\sum_{j=0}^{N} \iota_{ij} = 0, \qquad i = 1, \dots, n-1. \tag{7}$$

Problem 3 (regularized least squares): 4 points. Let $A \in \mathbb{R}^{m \times n}$. Consider the least squares problem:

$$\text{minimize} \quad \|Ax - b\|_2^2 \,. \tag{8}$$

When A is ill-conditioned, a regularized least squares problem can be solved instead, to improve the quality of the solution:

$$\text{minimize} \quad \|Ax - b\|_2^2 + \sigma \, \|x\|_2^2 \,, \tag{9}$$

where $\sigma > 0$. Compute the exact solution of this problem, and explain how it generalizes the solution of (8).

Problem 4 (Laplace's equation): 5 points. Let Ω be a compact subset of $\mathbb{R}^2$. Let $\partial\Omega$ be the boundary of Ω, and let $f : \partial\Omega \to \mathbb{R}$ be continuous. Laplace's equation with Dirichlet boundary conditions is:

$$\text{Find } u \text{ such that:} \quad \begin{cases} \Delta u(x) = 0, & x \in \operatorname{int}(\Omega), \\ u(x) = f(x), & x \in \partial\Omega, \end{cases} \tag{10}$$

where $\operatorname{int}(\Omega)$ denotes the interior of Ω. The solution u of Laplace's equation can be found by solving the following minimization problem:

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2} \int_\Omega \|\nabla u(x)\|^2 dx. \\ \text{subject to} \quad & u(x) = f(x), \quad x \in \partial\Omega. \end{aligned} \tag{11}$$

The integral here is called the Dirichlet energy.

1. Let $\Omega = [0,1] \times [0,1]$. Come up with a way of approximating the Dirichlet energy in order to turn this minimization problem into a finite-dimensional optimization problem.

2. Explain how to apply gradient descent to solve your problem.

Problem 5 (bonus problem): 2 points max. Program your algorithm from Problem 4 and see if you can get it to work. Include your code and some plots giving support.