
Department of Scientific Computing
Written Preliminary Examination
Spring 2012

January 6, 2012

Instructions:

- Solve only 10 of the 11 questions as completely as you can.
- All questions are weighted equally.
- All parts of a question are weighted equally unless stated otherwise.
- If you use web sources, please list them clearly.
- The exam is due back to Maribel Amwake no later than 1:00 pm on Monday, January 9, 2012; no exceptions allowed.
- If you have any questions related to this exam as you work on it, please send an e-mail to the person responsible, *and* Dr. Sachin Shanbhag (sshanbhag at fsu dot edu). The person responsible is listed at the beginning of each question.
- Write your Student ID on each of your answer sheets. Do not write your name on your answer sheets. When turning in your exam, include a cover page with your name and Student ID.

---

1. *Optimization: Augmented Lagrangian* (Dr. Navon)

Consider the equality constrained problem

$$\min f(x), \text{ subject to } g(x) = 0,$$

where $f : \mathbb{R}^n \to \mathbb{R}$, and $g : \mathbb{R}^n \to \mathbb{R}^m$ are continuously differentiable. The first order multiplier method finds,

$$x^k = \underset{x \in \mathbb{R}^n}{\text{argmin}}\ L_{\mu^k}(x, \lambda^k) \equiv f(x) + \lambda^{k'} g(x) + \frac{\mu^k}{2}||g(x)||^2.$$

and updates $\lambda^{k+1} = \lambda^k + \mu^k g(x^k)$. It increases the penalty coefficient μ^k such that $\mu^{k+1} \geq \mu^k$.
The algorithm of Augmented Lagrangian method is as follows:

1. Values of x_0, λ_0 and μ_0 are chosen to initialize the method. Then for $k = 0, 1, \cdots$, carry out optimality test: if $\nabla L(x_k, \lambda_k) = 0$ then stop.
2. Unconstrained minimization subproblem: Using any unconstrained minimization method you know solve:
$$\min_x L(x, \lambda_k, \mu_k) = f(x) - \lambda_k^T g(x) + \frac{1}{2}\mu_k g(x)^T g(x)$$
3. Update: Determine λ_{k+1} and μ_{k+1}.
$$\lambda_{k+1} = \lambda_k - \mu_k g(x_k + 1).$$
The penalty parameter should be chosen so that
$$\mu_k + 1 \geq \mu_k$$
This parameter should be large enough so that the augmented Lagrangian function has a local minimizer in x. If a failure is detected when attempting to solve this subproblem then penalty parameter μ_{k+1} should be increased.

Optimization Problem

$$\text{Minimize } f(x_1, x_2) = e^{3x_1} + e^{-4x_2} \text{ subject to } g(x_1, x_2) = x_1^2 + x_2^2 - 1 = 0$$

using the Augmented Lagrangian method. The Augmented Lagrangian function is:

$$L(x, \lambda, \mu) = e^{3x_1} + e^{-4x_2} - \lambda(x_1^2 + x_2^2 - 1) + \frac{1}{2}\mu(x_1^2 + x_2^2 - 1)$$

We start at initial guess $x_0 = (-1,\ 1)^T$, together with λ_0 = -1. For this example we will keep the penalty parameter constant $\mu_0 = \mu_k = 10$. At initial point we have:

$$\nabla f = \begin{bmatrix} 3e^{3x_1} \\ -4e^{-4x_2} \end{bmatrix} = \begin{bmatrix} 0.14936 \\ -0.07326 \end{bmatrix}, \qquad \nabla g = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$$

and,

$$\nabla^2 f = \begin{bmatrix} 9e^{3x_1} & 0 \\ 0 & 16e^{-4x_2} \end{bmatrix} = \begin{bmatrix} 0.44808 & 0 \\ 0 & 0.29305 \end{bmatrix}, \qquad \nabla^2 g = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

We use Newtons method for the unconstrained problem. For simplicity we use classical Newton method without a line search, that is: $x \leftarrow x - (\nabla^2_{xx} L)^{-1}(\nabla_x L)$.
For $x = x_0$,

$$\nabla_x L = \begin{bmatrix} -21.851 \\ 21.927 \end{bmatrix}, \qquad \nabla^2_{xx} L = \begin{bmatrix} 62.488 & -40.00 \\ -40.00 & 62.293 \end{bmatrix}$$

At this point, $x = [-0.78862, 0.78374]^T$, and $||\nabla_x L|| = 7.1533$. Continue iterating with Newtons method until $||\nabla_x L|| \leq 10^{-9}$.

Display only last iterates that satisfies above condition that is: $x_1 = [-0.71795, 0.63937]^T$. To complete the iteration of the Augmented Lagrangian method we update the Lagrange multiplier estimate:

$$\lambda_1 = \lambda_0 - \mu_0 g(x_1) = -1 - 10(-0.075757) = -0.24243$$

Now the whole process repeats at the new point. Complete writing the code for whole process that requires 6 full iterations of the Augmented Lagrangian method yielding a very accurate solution of: $x = [-0.74834, 0.66332]^T$, with $\lambda^* = -0.21233$.

You can use books and Matlab to write the Augmented Lagrangian code for solving this problem.

---

2. *Ordinary Differential Equations/Molecular Dynamics* (Dr. Shanbhag)

Consider the equations of motion of a single particle in one dimension, where $r(t)$, $v(t)$, and $F(t)$ denote the position, velocity and force on the particle.

$$\begin{aligned} \frac{dr}{dt} &= v \\ \frac{dv}{dt} &= \frac{F}{m} \end{aligned}$$

The Verlet algorithm considers a Taylor expansion of $r(t)$ as:

$$r(t+\Delta t) = r(t) + v(t)\Delta t + \frac{F(t)}{2m}\Delta t^2 + \frac{d^3 r}{dt^3}\frac{\Delta t^3}{6} + O(\Delta t^4)$$

which leads to an integration scheme that is fourth order in position, and second ordered in time.

$$r(t+\Delta t) + r(t-\Delta t) = 2r(t) + \frac{F(t)}{m}\Delta t^2 + O(\Delta t^4)$$

$$v(t) = \frac{r(t+\Delta t) - r(t-\Delta t)}{2\Delta t} + O(\Delta t^2)$$

Now consider the following scheme which is fourth order in both position and time.

$$\begin{aligned} r(t+\Delta t) &= r(t) + v(t)\Delta t + \frac{4f(t) - f(t-\Delta t)}{6m}\Delta t^2 \\ v(t+\Delta t) &= v(t) + \frac{2f(t+\Delta t) + 5f(t) - f(t-\Delta t)}{6m}\Delta t^2 \end{aligned}$$

- Show that the equation for position update is the same as the Verlet algorithm for the given choice of $v(t+\Delta t)$. (50 points)
- How would you verify whether the equation for velocity update is fourth ordered? It is sufficient to describe your method, without actually carrying it out. (25 points)
- In a real molecular dynamics implementation, a small memory footprint is desirable. Naively (in the equations above), we need to maintain two copies of position and velocity (at t and $t+\Delta t$), and three copies of force (at $t-\Delta t$, t and $t+\Delta t$). Suggest a way to implement the scheme that requires the maintenance of the least number of copies (that you can think of)? (25 points)

3. *Interpolation* (Dr. Shanbhag)

The population of the US is given in the following table:

Year	**Population**
1900	76,212,168
1910	92,228,496
1920	106,021,537
1930	123,202,624
1940	132,164,569
1950	151,325,798
1960	179,323,175
1970	203,302,031
1980	226,542,199

There is a unique polynomial of degree 8 that interpolates these 9 data points, but that polynomial can be represented in many different ways. Consider the following possible sets of basis functions $\phi_j(t),\ j = 1, ..., 9$: (i) $\phi_j(t) = t^{j-1}$, (ii) $\phi_j(t) = (t - 1900)^{j-1}$, (iii) $\phi_j(t) = (t - 1940)^{j-1}$, and (iv) $\phi_j(t) = ((t - 1940)/40)^{j-1}$.

- For the four sets of basis functions generate the corresponding Vandermonde matrix, and compute the corresponding condition number (use of a library routine is fine)? Explain your observations.
- Using the best-conditioned basis, compute the polynomial interpolant.
- Use Hermite cubic interpolation (again, use of a library routine is fine) to interpolate the data. Plot the data and both the interpolants.
- Extrapolate the population to 1990 using the two interpolants. How do they compare with the true value of 248,709,873 according to the 1990 census.

---

4. *Finite Element* (Dr. Burkardt)

Suppose that $u \in H^1(\Omega)$ is a solution of the Poisson equation $-\Delta u = f$ in the domain Ω, and that for some constant $\alpha > 0$, u satisfies the mixed boundary condition $\alpha u + \frac{\partial u}{\partial n} = 0$ on $\partial\Omega$.

Recall that $H^1(\Omega)$ is the set of functions $v : \Omega \to \mathbb{R}$ such that v and all first derivatives of v are square-integrable over Ω.

(a) Show that u satisfies the weak equation:

$$\int_\Omega \nabla u \cdot \nabla v + \alpha \int_{\partial\Omega} u\, v = \int_\Omega f\, v \quad \text{for all } v \in H^1(\Omega)$$

(b) For any $u, v \in H^1(\Omega)$, define:

$$B(u,v) \equiv \int_\Omega \nabla u \cdot \nabla v + \alpha \int_{\partial\Omega} u\, v$$

Show that $B(u,v)$ is an inner product on $H^1(\Omega)$.

(c) Use your answer to (b) to show that a solution of the weak equation is unique.

---

5. *Linear Algebra* (Dr. Burkardt)

Let A be a given $m \times n$ matrix, where $m > n$, and suppose A has rank n; let $\vec{b}$ be a given m-vector. The standard linear least squares problem for $A\vec{x} = \vec{b}$ is to find a vector which minimizes $\rho^2(x) = \|r\|_2^2$ where $\vec{r}$ represents the residual $\vec{r} = \vec{b} - A\vec{x}$ and $\|\cdot\|_2$ represents the standard ℓ_2 vector norm. We want to consider a generalization of this problem. With A defined as above, let B be a given symmetric positive definite $m \times m$ matrix. Consider the problem of minimizing

$$(\dagger) \qquad \sigma^2(x) = \|r\|_2 2 \quad \text{over all } \vec{x} \in \mathrm{R}^n$$

where now, instead of r being defined by the relationship $\vec{b} = A\vec{x} + \vec{r}$, r is instead implicitly defined by the relationship $\vec{b} = A\vec{x} + B\vec{r}$. Clearly, if $B = I$ Problem (†) reduces to the standard linear least squares problem.

a.) Show that (†) always has a solution.

b.) Show that $\vec{r} \in \mathcal{N}\big((B^{-1}A)^T\big)$; be sure to justify why $B^{-1}A$ is defined.

c.) Derive the analogue of the normal equations for this problem and explain how you would efficiently solve them.

d.) Suppose we have the following data:

t_i	1	2	4	5	10	16
f_i	6	1	2	3	4	5

Find the line which best fits the data in the least squares sense; for our generalized problem, this would imply $B = I$. Now modify $B = I$ by setting $B(1,1) = 100$ and keeping the remaining portion of B set to the identity. Find the line which solves (†) with this new B. Display your data points and the two approximating lines on one plot. Discuss your results and the implication of using a choice of $B \neq I$.

---

6. *Datastructures* (Dr. Beerli)

In phylogenetic research a key interest is finding the best relationship among species (finding the best tree). Define a datastructure that allows searching through all different such relationships with minimal operations.

- Give an example of the structure for multifurcating trees. These are relationship trees where more than two individuals can have the same ancestor. (20%)
- Give at least one algorithm to change from one multifurcating tree to another. (25%)
- Consider the following optimality criterion for a particular tree. Each species has a character that is measured the interval $(-\infty, \infty)$, for simplicity we assume that each branchlength on the tree is 1.0, the optimality score for the whole tree is a weighted average of the values on the tips (leaves) trickled down to the root of the tree, for example a tree with 3 species that have values a=1.3, b=1.6, and c=8.0 with a tree that combines a with b and then the ancestor ab with c, calculates as mean(mean(1.3,1.6),8.0)=4.725, but a tree with ((ac),b) results in 3.125, we will pick that tree with the highest score as the best fitting tree.

  (a) Present an algorithm to calculate this optimality score for any tree, in particular consider multifurcating trees. (35%)

  (b) After a rearrangement we need to recalculate the tree score, on large trees this is leads to many unnecessary calculations, present a way to minimize the calculations after a rearrangement. You may consider that the root is arbitrary, and can be moved for calculations. I only expect pseudo code here. (10%)

  (c) Present the optimality score for the **best** tree using the data below. (10%)

You can present the structure, and algorithms in pseudocode or a real programming language, but no phylogenetic packages/subroutines/modules are allowed)

Example dataset:

Toad	1.3
Frog	1.3
Newt	1.3
Snake	5.0
Monkey	10.0
Baboon	11.0
Fish	-4.0
Whale	8.0

---

7. *Linear Algebra* (Dr. Wang)

The matrix $C = (A^T A)^{-1}$, where $rank(A) = n$, arises in many statistical applications and is known as the *variance-covariance matrix*. In this problem, we try to build an algorithm to calculate the diagonal of C. Assume that the factorization

$$A = QR$$

is available.

1. Show that $C = (R^T R)^{-1}$.
2. Show that if we write
$$R = \begin{pmatrix} \alpha & v^T \\ 0 & S \end{pmatrix}$$
then
$$C = (R^T R)^{-1} = \begin{pmatrix} (1 + v^T C_1 v)/\alpha^2 & -v^T C_1/\alpha \\ -C_1 v/\alpha & C_1 \end{pmatrix}$$
where $C_1 = (S^T S)^{-1}$.
3. Using (2), give an algorithm that overwrites the upper triangular portion of R with the upper triangular potion of C.
4. Show how many flops your algorithm requires.

---

8. *Parallel Programming* (Dr. Wang)

Suppose the matrix L is a $n \times n$ lower triangle matrix, b is an n dimensional vector.

1. Write a Fortran, or C++ function to solve $Lx = b$.
2. Try to convert your code to a parallel OpenMP program by inserting some directives.
3. If (2) is impossible, please explain the reason. And please try to reorganize your code so that you can use OpenMP to parallelize it.

---

9. *Fourier Analysis* (Dr. Meyer-Baese)

I. A periodic signal g(t) is expressed by the following Fourier series:

$$g(t) = 2\cos 2t + 3\cos(3t - \pi/2) + 5\sin(6t - \pi/3) \tag{1}$$

a. Sketch the exponential Fourier series spectra.

b. Write the exponential Fourier series for $g(t)$.

II. From the definition $G(\omega) = \int_{-\infty}^{\infty} g(t)\exp^{-j\omega t}\, dt$ show that the Fourier transform of rect (t-5) is sinc$(\omega/2)e^{-j5\omega}$.

Hint: rect is given as

$$rect(x) = \begin{cases} 0 & : \quad |x| > 0.5 \\ 0.5 & : \quad |x| = 0.5 \\ 1 & : \quad |x| < 0.5 \end{cases} \tag{2}$$

and sinc(x) is given as sinc(x)=$\frac{\sin x}{x}$

10. *Partial Differential Equations* (Dr. Ye)

Heat flow in the earth is governed by the processes of conduction and convection. In regions where water is free to move, heat flow in the near surface (the top several hundred meters of the earths surface) is strongly affected by convection and the analysis of temperature change is quite complicated. In the Arctic, however, permafrost essentially renders water motion meaningless as a heat-flow mechanism. In these areas, conduction is the primary mechanism by which heat is transported in the crust and a relatively simple analysis may be appropriate. The equation for such a problem is:

$$\frac{\partial T}{\partial t} = \frac{K}{\rho c} \frac{\partial^2 T}{\partial z^2}$$

where T is temperature, z is depth below the surface, t is time, K is thermal conductivity, c is heat capacity, and ρ is density. Consider heat flow in the top 1,000m of the crust. The surface boundary condition is a specified temperature. The bottom boundary condition (at 1km) is that the upward heat flux, q, equal to $K\Gamma_0$, where Γ_0 is the geothermal gradient, about 3°C per 100m. The thermal conductivity of rock and of permafrost is about 0.5 $cal\ m^{-1}s^{-1\circ}\ C^{-1}$ and ρc is about 0.5 $cal\ cm^{-3}\ {}^{\circ}C^{-1}$.

1. Write a code to solve the temperature problem for permafrost regions. Explain how you determine the grid spacing.
2. Use your code to calculate the steady-state temperature profile for a surface temperature of $-15^{\circ}C$. Start the computations with $T = 0^{\circ}C$ everywhere.
3. Starting with the steady-state temperature profile as the initial condition, calculate the temperature profile every decade under global warming conditions of a steadily increasing surface temperature at a rate of $3.5^{\circ}C$ per century.

11. *Statistics* (Dr. Ye)

For a variable X, both confidence and credible intervals can be defined symbolically as $\text{Prob}(l \leq X \leq u) = 1 - \alpha$, where l and u are the lower and upper interval limits and α is significance level. However, the definition is interpreted in different ways when estimating regression confidence intervals and Bayesian credible intervals. Answer the following questions:

1. What are the conceptual differences between the two kinds of intervals?
2. Describe how the two kinds of intervals may be calculated.
3. Under what circumstances can one expect the two kinds of intervals to be the same?
4. The two intervals are always different. What are possible reasons that render the two kinds of intervals different?

For your convenience, you may limit your discussion in the context of parameter estimation and model predictions.

12. *Integration* (Dr. Beerli)

Numerically integrate

$$\int_{-5}^{5} \frac{\sin x3}{x} dx$$

- Solve the integral using a composite midpoint rule. Show the improvement of the result for several number of subintervals. (30%)
- Solve the integral using an iterative approach, Romberg integration. Show the improvement the results for several steps. (30%)
- Solve the integral using Monte Carlo, give a table with the number of trials and improvement of accuracy of the result. (30%)
- Compare the effort and error for all three methods. (10%)