Math 210C: Mathematical Physics

Instructor: Michael Holst

Spring Quarter 2016

Homework Assignment #1

Due Date: NONE (just some suggested problems to look at to complement the lectures)

Exercise 1.1. (*A Little Practice with Calculus in Banach Spaces*) Let X and Y be real Hilbert spaces, let $A \in \mathcal{L}(X,Y)$, $f \in Y$, and define $J: X \to \mathbb{R}$ as:

$$J(u) = \frac{1}{2}\|f - Au\|_X^2.$$

Show that a necessary condition for $J(u)$ to have a local minimizer $u \in X$ is

$$0 = J'(u) = A^T Au - A^T f,$$

where $A^T \in \mathcal{L}(Y,X)$ is the Hilbert-adjoint of A. *Hint:* Compute the Gateaux derivative $J'(u)$ of $J(u)$, and enforce the first-order necessary condition by setting $J'(u)$ to zero.

Exercise 1.2. (*A Little More Practice with Calculus in Banach Spaces*) Let X be a real Hilbert space, let $A \in \mathcal{L}(X,X)$, $f \in Y$, and define $J: X \to \mathbb{R}$ as:

$$J(u) = \frac{1}{2}(Au,u) - (f,u).$$

Show that a necessary condition for $J(u)$ to have a local minimizer $u \in X$ is

$$0 = J'(u) = \frac{1}{2}(A + A^T)u - f,$$

where $A^T \in \mathcal{L}(X,X)$ is the Hilbert-adjoint of A. Write a simpler form of the condition when A is self-adjoint. *Hint:* Compute the Gateaux derivative $J'(u)$ of $J(u)$, and enforce the first-order necessary condition by setting $J'(u)$ to zero.

Exercise 1.3. (*Taylor Expansion in Banach Spaces and Newton's Method*) Let X be a real Hilbert space, let $F: X \to X$, and consider the problem:

$$\text{Find } u \in X \text{ such that } F(u) = 0 \in X.$$

Using Taylor expansion, derive a simple Newton method (with uniform step-size $\alpha = 1$) for solving this problem:

- Choose arbitrary $u^0 \in X$, choose $TOL << 1$.
- For $k = 0, 1, 2, \ldots$ until ($\|F(u^k)\|_X < TOL$) do:
  1. Solve: $F'(u^k)p^k = -F(u^k)$,
  2. Update: $u^{k+1} = u^k + \alpha p^k$.
- End For.

Exercise 1.4. (*Directions of Decrease and Descent Directions in Optimization*) Let X be a real Hilbert space and let $J: X \to \mathbb{R}$. Recall that a *direction of decrease* $p \in X$ for J at the point $u \in X$ satisfies

$$J(u + \alpha p) < J(u), \quad \forall \alpha \in (0, \sigma], \text{ with } \sigma \text{ sufficiently small.}$$

Recall that a *descent direction* $p \in X$ for a functional $J: X \to \mathbb{R}$ at the point $u \in X$ satisfies

$$(J'(u), p) < 0.$$

Use a simple one-dimensional Taylor expansion to show that a descent direction p at u is always a direction of decrease p at u.

Exercise 1.5. (*The (Exact) Newton Direction as a Descent Direction*) Let X be a real Hilbert space, let $F\colon X \to X$, and consider again the problem:

$$\text{Find } u \in X \text{ such that } F(u) = 0 \in X.$$

One way to make Newton's method more robust for this problem is to construct an associated minimization problem such that the minimum occurs at the solution to the original nonlinear equation. A specially engineered functional to minimize that accomplishes this for Newton's method is the following:

$$J_F(u) = \frac{1}{2}\|F(u)\|_X^2 = \frac{1}{2}((F(u), F(u))_X.$$

This is a real-valued function of a function, and we can form its derivatives, as long as $F(u)$ is differentiable.

1. Your first task is now to show that the (exact/full) Newton direction $p = -F'(u)^{-1}F(u)$ is actually a descent direction (hence a direction of decrease) for this specially engineered functional $J\colon X \to \mathbb{R}$. To show this, you will need to compute the Gateaux derivative of $J_F(u)$ (show your work):

$$J_F'(u) = F'(u)^T F(u),$$

   where $F'(u)$ is the Gateaux derivative of $F(u)$ with respect to u, and $F'(u)^T$ is its adjoint as a linear operator on X.
2. Once you have derived $J_F'(u)$, modify the Newton method you derived earlier to produce a *Damped Newton Iteration* that generates iterates u^k that are guaranteed (through control of steplength α) to reduce the value of the special functional J_F at each step. (This is sometimes called "globalizing Newton iteration"; the algorithm you have produced is at the core of nearly all of computational science.)

Exercise 1.6. (*The In-Exact Newton Direction as a Descent Direction*) Let X be a real Hilbert space, let $F\colon X \to X$, and consider again the problem:

$$\text{Find } u \in X \text{ such that } F(u) = 0 \in X.$$

Often the Newton equations are not solved exactly for reasons of efficiency (e.g., you are left with some residual $r = -F(u) - F'(u)w$). How large can one allow r to get and still end up with a descent direction for $J_F(u)$? Use your results from the earlier problems to derive a necessary and sufficient condition on r for descent, and then develop a somewhat simpler sufficient condition.

Exercise 1.7. (*Minimization with Equality Constraints and Implicit Manifolds*) Let X and Y be a real Banach spaces, let $f\colon X \to \mathbb{R}$, and let $g\colon X \to Y$. Consider the equality-constrained optimization problem:

$$\min_{u \in X} f(u),$$
$$\text{s.t. } g(u) = 0.$$

If we define the set of "feasible points" $u \in X$ that satisfy the constraint $g(u) = 0$:

$$M = \{\ u \in X \mid g(u) = 0\ \},$$

then our optimization problem above can also be written as:

$$\min_{u \in M \subseteq X} f(u).$$

Our first observation was that a necessary condition for f to have a (local) minimizer $u \in M$ is that the (Frechet) derivative of f at u is zero in all directions in the tangent space $T_u M$ to M at u:

$$\langle f'(u), v\rangle = 0, \forall v \in T_u M.$$

We also noted that characterizing $T_u M$ seems quite implausible, so this does not seem helpful. However, if a key mathematical relationship holds, than it allows us to characterize $T_u M$ completely, using only things we can compute. That relationship is: $T_u M = \mathcal{N}(g'(u))$, i.e., the tangent space to M at u is simply the null space of the linear operator $g'(u)\colon X \to L(X,Y)$ that arises as the Frechet derivative of g evaluated at u. When this relationship holds, we have access to the following general theorem about *Lagrange multipliers/functionals*:

Theorem 0.1. *If g is a submersion at $u \in M$, then f is stationary at u with respect to M if and only if the Euler-Lagrange equations hold for a fixed "Lagrange" functional $\lambda \in Y^*$:*

$$\langle f'(u), v\rangle - \langle \lambda, g'(u)v\rangle = 0, \qquad \forall v \in X, \tag{1.1}$$
$$\langle \eta, g(u)\rangle = 0, \qquad \forall \eta \in Y^*. \tag{1.2}$$

The key value in this theorem is that we now consider all directions $v \in X$, and $T_u M$ no longer appears anywhere. This now looks very similar to the first order necessary condition for optimality when there was no constraint function $g(u)$. Lastly, recall that the property that g be a submersion at $u \in X$ was simply that:

1. $g(u)$ is C^1 in a ball around u
2. $g'(u)\colon X \to L(X,Y)$ is a surjective (onto) linear operator at u (i.e., $\mathcal{R}(G'(u)) = Y$).
3. X can be split into $\mathcal{N}(G'(u))$ and $\mathcal{N}^{\perp}(G'(u))$.

When g has these properties, one can show that M has the structure of a differentiable manifold (and comes with all of the things that manifolds have, such as an atlas of charts, etc), and this is what makes the method of Lagrange multipliers work. Lastly, we noted that when $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$, and typically with $0 < m < n$, then there are some simple and very easily verified sufficent conditions for g to be a submersion. (For example, if $g'(u)$ is a full-rank matrix when evaluated at u, then g is a submersion.)

Finally (after that long review/summary), for this first problem, show that the Euler-Lagrange equations (**??**)–(**??**) arise by simply setting to zero the Gateaux derivative of a "Lagrangian" with respect to the variable u (giving (**??**)) and with respect to the variable λ (giving (**??**)), where the Lagrangian $L\colon X \times Y^* \to \mathbb{R}$ is defined as:

$$L(u, \lambda) = f(u) - \langle \lambda, g(u) \rangle. \tag{1.3}$$

Exercise 1.8. (*Equality Constraints and Implicit Manifolds*)

Let X and Y be a real Banach spaces, let $f\colon X \to \mathbb{R}$, and let $g\colon X \to Y$. Consider again the equality-constrained optimization problem:

$$\begin{aligned} &\min_{u \in X} f(u), \\ &\text{s.t. } g(u) = 0. \end{aligned}$$

Now, however, let us assume that X and Y have the additional structure of a Hilbert space (they have inner-products that induce their norms), and let us assume that f and g have particularly simple forms:

$$f(u) = \frac{1}{2}(Au, u)_X - (b, u)_X, \qquad g(u) = c - Bu, \tag{1.4}$$

where $b \in X$, $c \in Y$, $A \in L(X,X)$, and $B \in L(X,Y)$. In addition, assume that A is self-adjoint ($A = A^T$). Show that the first-order necessary condition for our equality-constrained problem to have a solution is the following block linear system of equations:

$$\begin{array}{ccccc} Au & + & B^T\lambda & = & b, \\ Bu & & & = & c. \end{array} \tag{1.5}$$

Exercise 1.9. (*Equality Constraints and Implicit Manifolds*)

A pendulum in the plane is a classical example of a mechanical system with a (holonomic) constraint. The planar pendulum is characterized by the position $q_i(t) = (q_1(t), q_2(t))$ and velocity $\dot{q}_i(t) = (\dot{q}_1(t), \dot{q}_2(t))$ of the pendulum "bob", modeled as a point with mass m, and the fixed length l of the pendulum rod. The kinetic energy, potential energy, and constraint of the mechanical system are respectively

$$T(\dot{q}_i) = \frac{m}{2}(\dot{q}_1^2 + \dot{q}_2^2), \qquad V(q_i) = mgq_2, \qquad G(q_i) = (q_1^2 + q_2^2) - l^2 = 0, \tag{1.6}$$

where g is the acceleration due to gravity. The constraint is simply that the pendulum bob must always be a distance l from the center of the pendulum (taken to be at the origin for simplicity). The Lagrangian and Hamiltonian then have the form

$$L(q_i, \dot{q}_i) = T(\dot{q}_i) - V(q_i) = \frac{m}{2}(\dot{q}_1^2 + \dot{q}_2^2) - mgq_2, \tag{1.7}$$

$$H(q_i, p_i) = T(p_i) + V(q_i) = \frac{1}{2m}(p_1^2 + p_2^2) + mgq_2, \tag{1.8}$$

where we have employed the Legendre transformation $p_i = \partial L / \partial \dot{q}_i = m\dot{q}_i$.

1. Assume the rod breaks at time zero, so we can ignore constraint. Derive the equations of motion for the pendulum bob from the action integral using the Lagrangian framework. I.e., don't just write down the Euler-Lagrange equations, but instead actually derive them from the action integral

$$S(q_i) = \int_0^T L(q_i, \dot{q}_i)\ dt,$$

by taking its Gateaux derivative with respect to q_i and setting it to zero.

2. Now do the same thing (assume rod breaks at time zero) in the Hamiltonian framework, by starting instead with this action integral:

$$S(q_i, p_i) = \int_0^T \sum_{j=1}^{n} p_j \dot{q}_j - H(q_i, p_i)\ dt.$$

3. Now assume that the rod holds at all times, so we have to incorporate the single constraint into the equations of motion. We use our powerful Lagrange multiplier framework, and form the augmented Lagrangian:

$$\bar{L}(q_i, \dot{q}_i) \quad = \quad L(q_i, \dot{q}_i) - \lambda G(q_i) = \frac{m}{2}(\dot{q}_1^2 + \dot{q}_2^2) - mgq_2 - \lambda G(q_i). \tag{1.9}$$

Derive the Euler-Lagrange equations by taking the Gateaux derivative of the action integral built from this augmented Lagrangian as before, but now with respect to both q_i and λ_i, to produce:

$$\begin{bmatrix} m\ddot{q}_1 \\ m\ddot{q}_2 \\ 0 \end{bmatrix} = \begin{bmatrix} -2q_1\lambda \\ -mg - 2q_2\lambda \\ q_1^2 + q_2^2 - l^2 \end{bmatrix}, \tag{1.10}$$

which is a second-order (ordinary) differential algebraic equations (DAE) for the pendulum configuration over time.

4. Now do the same thing but in the Hamiltonian framework, assuming the rod holds and we need to incorporate the single constraint into the equations of motion. We again use our powerful Lagrange multiplier framework, and form the augmented Hamiltonian:

$$\bar{H}(q_i, p_i) \quad = \quad H(q_i, p_i) + \lambda G(q_i) = \frac{1}{2m}(p_1^2 + p_2^2) + mgq_2 + \lambda G(q_i). \tag{1.11}$$

Derive the Hamilton's equations by taking the Gateaux derivative of the action integral built from this augmented Hamiltonian as before, but now with respect to all of q_i, p_i, and λ_i, to produce:

$$\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{p}_1 \\ \dot{p}_2 \\ 0 \end{bmatrix} = \begin{bmatrix} m^{-1}p_1 \\ m^{-1}p_2 \\ -2q_1\lambda \\ -mg - 2q_2\lambda \\ q_1^2 + q_2^2 - l^2 \end{bmatrix}, \tag{1.12}$$

which is a first-order (ordinary) differential algebraic equations (DAE) for the pendulum configuration over time.