OPTIMISATION 2 EXAM PREPARATION GUIDELINES

This points out some important directions for your revision.

General:

- The exam is fully based on what was taught in class: lecture notes, handouts and homework. Some of the problems you may have seen, some are new but do not require any background beyond the taught content of the course. Past exams can be used as practice. The format of the exam is – your best 4 problems out of 5, done over 2.5 hours will count. Recall that you are allowed one double-sided crib sheet whereon you can put down anything you want, and a simple non-graphical calculator to do arithmetics if you wish. You can write anything on your crib, it's up to you.
- Proofs of long and difficult theorems, such as the Basic solutions theorem, Separating Hyperplane, Strong duality theorem (when it is derived from Farkas), KT theorem are non-examinable. Look at the past exams to identify the level of "proof-type" problems you may expect.

Fundamentals

Simple optimisation problems. You should understand the ideas underlying mathematical optimisation and how they relate to the geometric properties of the feasible set and the level sets of the objective function. These ideas get generalised and pushed into higher dimensions by using linear algebra in LP and differential calculus in NLP. Review the problems under the *General optimisation problems* heading in homework 1 as well as problem 1 in the past three years' exams. For many examples of optimisation problems see the class notes and/or Franklin's article on http://www.maa.org/pubs/Calc_articles/ma075.pdf (there is a link to it from the course web page). If there is a somewhat difficult problem of this kind, you amy expect to have seen you, so review all the examples that had been discussed.

The manufacturing (MP) and diet problem (DP)

1. MP

$$\text{primal : Max } \boldsymbol{c}\cdot\boldsymbol{x},\ \text{s.t. } A\boldsymbol{x}\le\boldsymbol{b},\ \boldsymbol{x}\ge 0;\ \ \text{dual : Min } \boldsymbol{b}\cdot\boldsymbol{y},\ \text{s.t. } A^T\boldsymbol{y}\ge\boldsymbol{c},\ \boldsymbol{y}\ge 0.$$

   Canonical form

$$\text{Max } \boldsymbol{c}\cdot\boldsymbol{x}+0\boldsymbol{s},\ \text{s.t. } A\boldsymbol{x}+I\boldsymbol{s}=\boldsymbol{b},\ \boldsymbol{x},\boldsymbol{s}\ge 0,$$

   where $\boldsymbol{s}$ are slack variables and I the identity matrix.

   Exercise: Show that the above primal-dual relation is compatible with the definition of the dual problem for the Canonical form (which as you should remember does not require $\boldsymbol{y}\ge 0$).

   Weak duality, know how to prove it: if $\boldsymbol{x}$ is feasible for the primal and $\boldsymbol{y}$ for the dual, then $\boldsymbol{c}\cdot\boldsymbol{x}\le\boldsymbol{b}\cdot\boldsymbol{y}$.

   Strong duality: Equality of the objectives occurs only if $\boldsymbol{x}$ and $\boldsymbol{y}$ are optimal for the primal/dual.

   Complementary slackness: Suppose $\boldsymbol{x},\boldsymbol{y}$ are optimal. If some MP constraint is satisfied as a strict inequality (the difference between the r.h.s and l.h.s being the excess amount) the constraint's shadow price, the corresponding component of $\boldsymbol{y}$, is zero.

   Complementary slackness: Suppose $\boldsymbol{x},\boldsymbol{y}$ are optimal. If a dual inequality is satisfied as a strict inequality, then the corresponding component of $\boldsymbol{x}$ is free, i.e. zero (economically - the corresponding good is not manufactured). The difference between the l.h.s. and the r.h.s for the dual inequalities is called the reduced cost of a variable. Note, if x is an optimizer for the primal, the reduced cost of its every basic (strictly positive) component is zero.

Optimality test for $\boldsymbol{x}$: check feasibility, and then use Complementary Slackness to find a feasible $\boldsymbol{y}$. If this is successful, i.e a feasible $\boldsymbol{y}$ satisfying CS has been found, then both $\boldsymbol{x}, \boldsymbol{y}$ are optimal.

Sensitivity: If optimal $\boldsymbol{x}, \boldsymbol{y}$ have been found, the value of the problem can be computed in both ways: as $\boldsymbol{c} \cdot \boldsymbol{x}$ and $\boldsymbol{b} \cdot \boldsymbol{y}$. Which means, for small variations of $\boldsymbol{b}$, we use the latter, and for smaller variations of $\boldsymbol{c}$ – the former expression.

As far as the changes of $\boldsymbol{b}$ are concerned, the basis for $\boldsymbol{x}$ remains optimal, as long as it remains feasible, because this does not affect our solution for $\boldsymbol{y}$ using complementary slackness ($\boldsymbol{b}$ does not enter the dual constraints).

But the variations of $\boldsymbol{c}$ may affect feasibility of $\boldsymbol{y}$. The won't if the variations of $\boldsymbol{c}$ are small enough, and we are not in the alternative solutions case. But if we are, then one of the bases for $\boldsymbol{x}$, as in HW 2 Problem 3(d) will stop being optimal. But only one of the two: the other shall stay optimal. Note that changes of $\boldsymbol{c}$ do not affect feasibility of $\boldsymbol{x}$, because $\boldsymbol{c}$ does not enter the constraints for $\boldsymbol{x}$. To see which one (of the two optimal bases) survives, it suffices to proceed as in Problem 3(d), and use Strong duality to calculate the change of value.

Review: Problems 1–3 in HW2 (a must!), and past exams.

2. Note: If you are to solve a DP, you may prefer to solve the dual MP and then find shadow prices. They will yield the optimal solution for the original DP.

Simplex method

1. Canonical form – for any LP.

2. Notion of the basic solution and basis. Theorem on basic solutions, without proof. Using it for graphical solutions of problems with two constraints, as in HW3 (see also Problem 1 in the 2009 exam).

3. Simplex tableau algorithm. Short or long tableaus – up to you. But never have negative numbers in the value column, $\boldsymbol{x}$-rows.

   Understand how the initial tableau is constructed, in terms of basic and free variables. Always be able to pass form a tableau to a systems of equations and backwards. Understand the choice of rows and columns for pivoting. Understand the role of the entries in the free columns in the objective row. Understand the role of the ratios, computed at each step of the procedure – they give you the values for the new BFS. Namely, having chosen a free variable which is to become basic, the minimum positive ratio simply gives you the new value for this variable when it will become basic in the ensuing tableau. Then the product of this ratio with the corresponding entry in the objective row gives you improvement in your objective after the chosen free variable will have been brought into the basis in the ensuing tableau.

   Be able to identify a final tableau and write out the solution it provides (*including the unbounded case and the alternative solutions case*), both for Max and Min problems.

4. Two-phase method. On Phase I one always tries to *minimise* the sum of artificial variables. Note: a tableau is good enough for SM only after the objective variable has been is expressed via free variables only.

   After Phase I is completed, know how to write out the BFS found or to identify the problem as unfeasible. Know how to modify a tableau used for Phase I to proceed with Phase II.

5. Dual simplex method: It's non-examinable in principle. However what you may need is that in the Manufacturing problem, the final tableau objective row entries in slack variables' columns yield shadow prices of the constraints. In general, just remember, no matter whether it is two-phase or not, you keep track of all columns that in the initial tableau created a unit matrix (Phase II – so not erase artificial variables' columns, but do not pivot them any more). The entries in the objective row in the final tableau is the shadow price of a constraint, corresponding to the original position of 1 in the initial tableau.

6. SM and duality. See formulas (3–9) in the homonymous handout. These formulas tell you how to set up a tableau, given the basis. What you really do, given the basis, is simply express the basic variables and the objective via free variables. Tableau-wise, you make the unit matrix out of the basic sub-matrix (but there is no recipe as to which column of the unit matrix is which, but trial and error ... trial would do). A must: review problem 2 in HW4 and similar past exam problems.

Review: HW 3,4 everything on SM. Play around a bit with the online solver: try to come up with a problem of your own and then solve it, first using the solver http://riot.ieor.berkeley.edu/riot/Applications/SimplexDemo/Simplex.html and then by hand. Look at past exams. Bear in mind: very ugly calculations might mean you're doing something wrong.

Linear programmes and their duals

Reread the two handouts on duality.

1. Be able to express the changes brought by casting a LP into canonical form in the matrix form, i.e., keep track of how the quantities $\boldsymbol{x}, A, \boldsymbol{c}, \boldsymbol{b}$ change.

2. Be able to write a dual to any LP. This is done by first via casting the original LP into canonical form, writing the dual, and then returning to the original formulation in terms of the original quantities $\boldsymbol{x}, A, \boldsymbol{c}, \boldsymbol{b}$. Review problem 2 HW2, and HW5, see also past exam papers.

   Some basic principles: shadow prices of $\leq$ $(\geq)$ constraints are positive (negative), while for $=$ constraints shadow prices are unsigned. Works the other way around as well: an unsigned variable would produce an equation, rather than inequality, in the dual constraint.

3. Optimality check in SM: Your tableau calculations, boil down in essence to the following. Given a BFS $\boldsymbol{x}$, we've split the quantities $(\boldsymbol{x}, A, \boldsymbol{c})$ in the *canonical form* into basic and free components, e.g. $\boldsymbol{x} = (\boldsymbol{x}_b, \boldsymbol{x}_f)$. Then we have, in fact, solved $A_b^T \boldsymbol{y} = \boldsymbol{c}_b$ for $\boldsymbol{y}$. Finally, if $\boldsymbol{y}$ you've found satisfies the rest of the dual inequalities (for free components of $\boldsymbol{c}$) then $(\boldsymbol{x}, \boldsymbol{y})$ are optimizers for the primal/dual pair. Whether or not this is the case ir reflected by reduced costs in the tableau. If they are all non-negative, for the Max problem, then $\boldsymbol{x}$ is optimal (and so is $\boldsymbol{y}$). Otherwise $\boldsymbol{x}$ is not optimal, and if you were doing a simplex method, you would have to do more pivoting. Understand how this fact is reflected in the simplex tableau computation.

   In other words, a basis is *optimal* for the primal if the corresponding $\boldsymbol{y}$ (obtained from the dual with basic dual constraints taken as equations) is *feasible* for the dual. (This also pertains to the above section about MP–DP.) The fact that these $\boldsymbol{x}$ and $\boldsymbol{y}$ alone does not suffice – this is ensured by the construction of $\boldsymbol{y}$.

4. Shadow prices and reduced costs: If $\boldsymbol{x}$ passes the optimality check, then the found $\boldsymbol{y}$ is a vector of shadow prices, and the difference between the l.h.s and the r.h.s. in the dual inequalities for

free components of $\boldsymbol{c}$ are reduced costs. Hence, for a basic component of $\boldsymbol{x}$, its reduced cost is always zero, by definition.

5. For MP: If a slack variable ends up being free (a constraint is satisfied without a slack) then its reduced cost of this variable and the shadow price of the corresponding constraint coincide.

Go back to HW3,4,5 and make sure you understand what is going on from the duality viewpoint on each step of the simplex method computations.

Geometry of LP, Extreme value theorem, and Farkas

What was taught about open and closed sets can be regarded as background material and the test will not focus on it. However, you are expected to

1. Understand what closed and convex sets are.To prove that a set is convex, you either use the definition, or claim that it is a sublevel set of a convex function.

   As for closed sets, stating that a set defined in terms of $\leq$, $\geq$, $=$ constraints is closed suffices.

2. Understand the formulation of the Farkas alternative, its geometric meaning and applications, such as problems 3–6 in HW6, the role of Farkas in the proof of the Kuhn-Tucker theorem. Look at the past exams and the corresponding Farkas problems. Look at its application to Markov matrices in the handout.

3. Understand, in the same way as in Duality how Farkas changes its statement when it is played with: e.g. if $A\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{x} \geq 0$ changes to $A\boldsymbol{x} \leq \boldsymbol{b}$, $\boldsymbol{x} \geq 0$, $\boldsymbol{y}$ acquires a sign constraint; removing $\boldsymbol{x} \geq 0$ on the primal side of Farkas changes the $A^T\boldsymbol{y} \geq 0$ on the dual side to $A^T\boldsymbol{y} = 0$ (making it Fredholdm alternative). Remember, all these changes come from the same basic tricks: introduction of slack variables or setting $\boldsymbol{x} = \boldsymbol{u} - \boldsymbol{v}$ for unsigned $\boldsymbol{x}$.

Game Theory

Understand what pure and mixed strategies and expected payoffs, in terms of the payoff matrix A are (if P plays $\boldsymbol{p}$ and Q plays $\boldsymbol{q}$, the expected payoff equals $\boldsymbol{q} \cdot A\boldsymbol{p}$.) Be able to solve simple games, given the payoff matrix, a la Problems 1,2 in HW 6, game section. (Tip – to solve the arising LPs you may wish to avoid using the simplex method! Also, sometimes, *but not always!* , some pure strategies can be discarded right away. E.g. if the payoff matrix has one column (1,2,3) and the other (2,3,4) then the former strategy shall never be played, because whatever the opponent does, playing the latter strategy is clearly more advantageous. Such little things enable you to save time and avoid extra calculations.)

Note: Problem 2 HW 6 was a 2010 exam problem, but the online version of the exam paper is old and gives a wrong answer to the problem (why?) see HW6 solutions for the right one.

Also, know how to punish: if Q plays a given non-optimal $\boldsymbol{q}$, P simply solves max $\boldsymbol{p} \cdot A^T\boldsymbol{q}$, under the only constraint $\sum p_j = 1$, and a pure strategy does it.

Unconstrained extrema, critical points, Hessian

Review the corresponding handout, homework and past exams and make your own conclusions.

1. Be able to find and characterise critical points of functions.

2. Distinguish *local* and *global* extrema (extremisers). Be able to visualise the feasible set and the level sets of the objective function and on this basis conjecture the answer to the problem.
3. Understand the simplifying role of convexity. I.e. if a function is convex (concave), i.e. the Hessian is positive (negative) definite *everywhere*, then a local minimum (maximum) is the global one.
4. Understand how Extreme value theorem can facilitate your judgements about local extrema being global.

Convex functions

Review the handout and the first batch of problems in homework 7 and past exams.

1. Know how to use definitions and convexity criteria, such as the second derivative test (you do not have to know how to derive it). Understand the notion of a function's sublevel set and how it connects convex functions with convex sets. Understand the generalisation of the convexity definition, as Jensen's inequality.
2. Recall that in order to prove that a certain set is convex, it may be convenient to show that it is a sublevel set of a convex function, and sublevel sets of convex functions are convex.
3. Inequalities: Jensen, Cauchy, and Cauchy-Schwartz - know how to derive and when to use them. Especially Cauchy-Schwartz: see past exams and homework. Review the Applications of Cauchy-Schwarz handout, at least the first two applications. (The last one may be a bit lengthy for the exam).

Equality constraints: Lagrange multipliers

Review the corresponding handout as well as the second bunch of problems in homework 7, and on the past exams.

1. Be able to formulate and solve problems. Before you start actually solving it, make sure that you have done everything you can to simplify it.

   There are no recipes to solving the arising systems of nonlinear algebraic equations. However, symmetry of the system of equations with respect to permutations of the variables often enables one to find solutions very efficiently (which happens in "real life" applications).
2. Distinguish *local* and *global* extrema (extremisers). Understand the role of convexity. Recall that the second-derivative test should apply to the Lagrangian, after fixing the value of the Lagrange multipliers. However, in "real-life" problems one is often better off with a general argument about the existence of a maximizer or minimizer, rather than doing an often unwieldy calculation, involving the Hessian. E.g. Extreme value theorem can facilitate your judgements. But if need be, the Hessian is computed w.r.t the $\boldsymbol{x}$ variables only, for *each* value of $\boldsymbol{\lambda}$, found after solving the Lagrange equations.
3. Be able to relate the Lagrange multipliers method to its generalisation in the KT theory. Observe that one can treat LM method as a special case of KT ($\boldsymbol{g} \geq 0$ and $\boldsymbol{g} \leq 0$). This will simply remove the requirement $\boldsymbol{\lambda} \geq 0$ from the KT theorem.

Kuhn-Tucker conditions

Review the very last handout and corresponding problems in HW8, and, similar to them, past exams. Draw your own conclusions on how imaginative I can be with KT problems.

Remember: the philosophy here is to attempt to describe the properties of a local extremiser of an optimisation problem, assuming that it has been found. Mathematically this description boils down to KT conditions. Not knowing an extremiser $\boldsymbol{x}$ and using KT, one is trying to go the other way around: find *all* the solutions of the KT conditions for a given optimisation problem (hopefully not too many), and then analyse them to decide on which of the solutions found yield the extremisers.

If there are no solutions, but the extremum should exist, because we are dealing with continuous functions on closed and bounded sets, then the extremiser must be one of those (hopefully few) points where the CQ non-degeneracy conditions are not satisfied. CQ may fail at points, where gradients of the constraints, tight at these points, become linearly dependent.

1. Be able to cast any optimisation problem into a suitable KT theorem input: $\mathsf{Max} f(\boldsymbol{x})$, s.t. $\boldsymbol{g}(\boldsymbol{x}) \geq 0$, $\boldsymbol{x} \geq 0$. Remember that $\boldsymbol{x} \geq 0$, if they are there, are constraints as well, only very simple, with constant gradients equal to $\boldsymbol{e}^j$, unit vectors in the coordinate directions.

2. Be able to identify constraints that matter, i.e. are tight at a given point $\boldsymbol{x}$, as well as the set of feasible directions $FD(\boldsymbol{x})$, in terms of their gradients. Understand the difference between the set $TFD(\boldsymbol{x})$ of *true feasible directions* – defined geometrically as the set of all initial directions of feasible curves starting at $\boldsymbol{x}$ – and the set of *feasible directions*, which consists of all vectors $\boldsymbol{v}$, such that $\boldsymbol{v} \cdot \nabla g(\boldsymbol{x}) \geq 0$ for each constraint g that matters at $\boldsymbol{x}$. Observe that one always has the inclusion $TFD(\boldsymbol{x}) \subseteq FD(\boldsymbol{x})$. The non-degeneracy CQ – constraint qualifications at $\boldsymbol{x}$ – means to show that any *feasible direction* is in fact a *true feasible direction.* I.e the above set inclusion is, in fact, equality. It is only when CQ are satisfied, that the KT theorem is actually a theorem.

3. Review how KT interact with the LP duality theory, see the KT handout.

Good luck, and make it easy for me!