Optimization, Quiz

May 8, 2008

- You will have one hour to complete this exam.
- Remember to write your name on this cover sheet! (below)

Name:

1

Which of the following are convex? Which of the following are concave? Check one box for each part:

1. $f(x) = \min_i x_i$ on $\mathbb{R}^n$ Convex: ☐ Concave: ☐ Neither: ☐
2. $f(x) = \max_i x_i$ on $\mathbb{R}^n$ Convex: ☐ Concave: ☐ Neither: ☐
3. $f(x) = \min_i |x_i|$ on $\mathbb{R}^n$ Convex: ☐ Concave: ☐ Neither: ☐
4. $f(x) = \max_i |x_i|$ on $\mathbb{R}^n$ Convex: ☐ Concave: ☐ Neither: ☐
5. $f(x) = \begin{cases} 1 & \text{if } |x| \leq 3 \\ 0 & \text{otherwise} \end{cases}$ on $\mathbb{R}$ Convex: ☐ Concave: ☐ Neither: ☐
6. $f(x) = \begin{cases} \infty & \text{if } |x| \leq 3 \\ 0 & \text{otherwise} \end{cases}$ on $\mathbb{R}$ Convex: ☐ Concave: ☐ Neither: ☐
7. $f(x) = \begin{cases} 0 & \text{if } |x| \geq 3 \\ 1 & \text{otherwise} \end{cases}$ on $\mathbb{R}$ Convex: ☐ Concave: ☐ Neither: ☐
8. $f(x) = \begin{cases} 0 & \text{if } |x| \geq 3 \\ -\infty & \text{otherwise} \end{cases}$ on $\mathbb{R}$ Convex: ☐ Concave: ☐ Neither: ☐

2

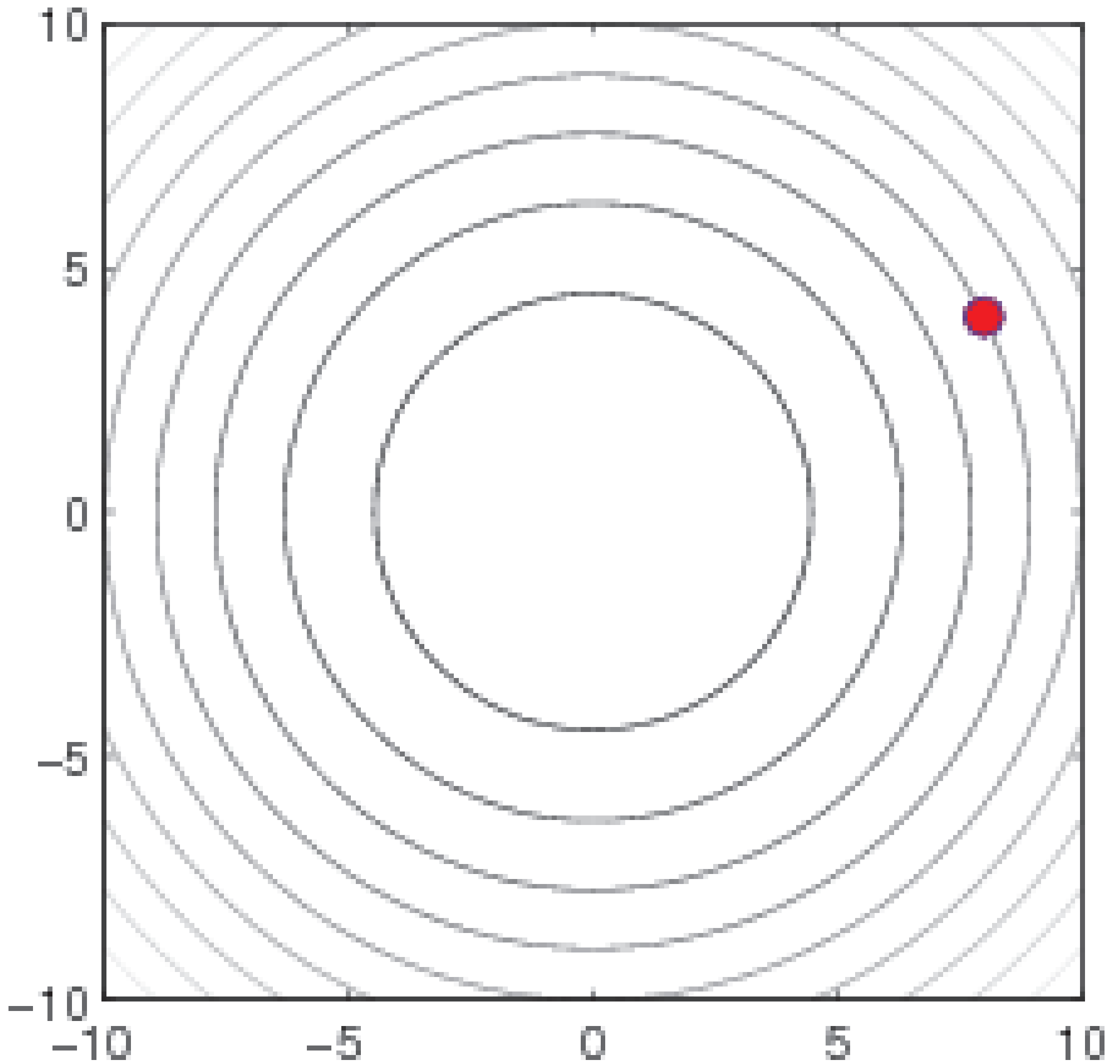
In this problem, we ask that you indicate the first three iterates of various optimization methods run on a quadratic objective. For each objective, each initial point $x^{(0)}$, and each method, indicate the approximate locations of $x^{(1)}$, $x^{(2)}$ and $x^{(3)}$ (unless the method terminates before then). You will do this by noting the points on the plot with letters, and then indicating for each method, what letters correspond to the first, second and third iterates. Of course, the same point might appear as an iterate for several methods. For example, you answer might look something like:

1. Method 1 $x^{(1)} = A \;\; x^{(2)} = B \;\; x^{(3)} = C$
2. Method 2 $x^{(1)} = A \;\; x^{(2)} = C \;\; x^{(3)} = D$
3. Method 2 $x^{(1)} = A \;\; x^{(2)} = B \;\; x^{(3)} = \text{done}$

With the four points A, B, C and D indicated on the plot.

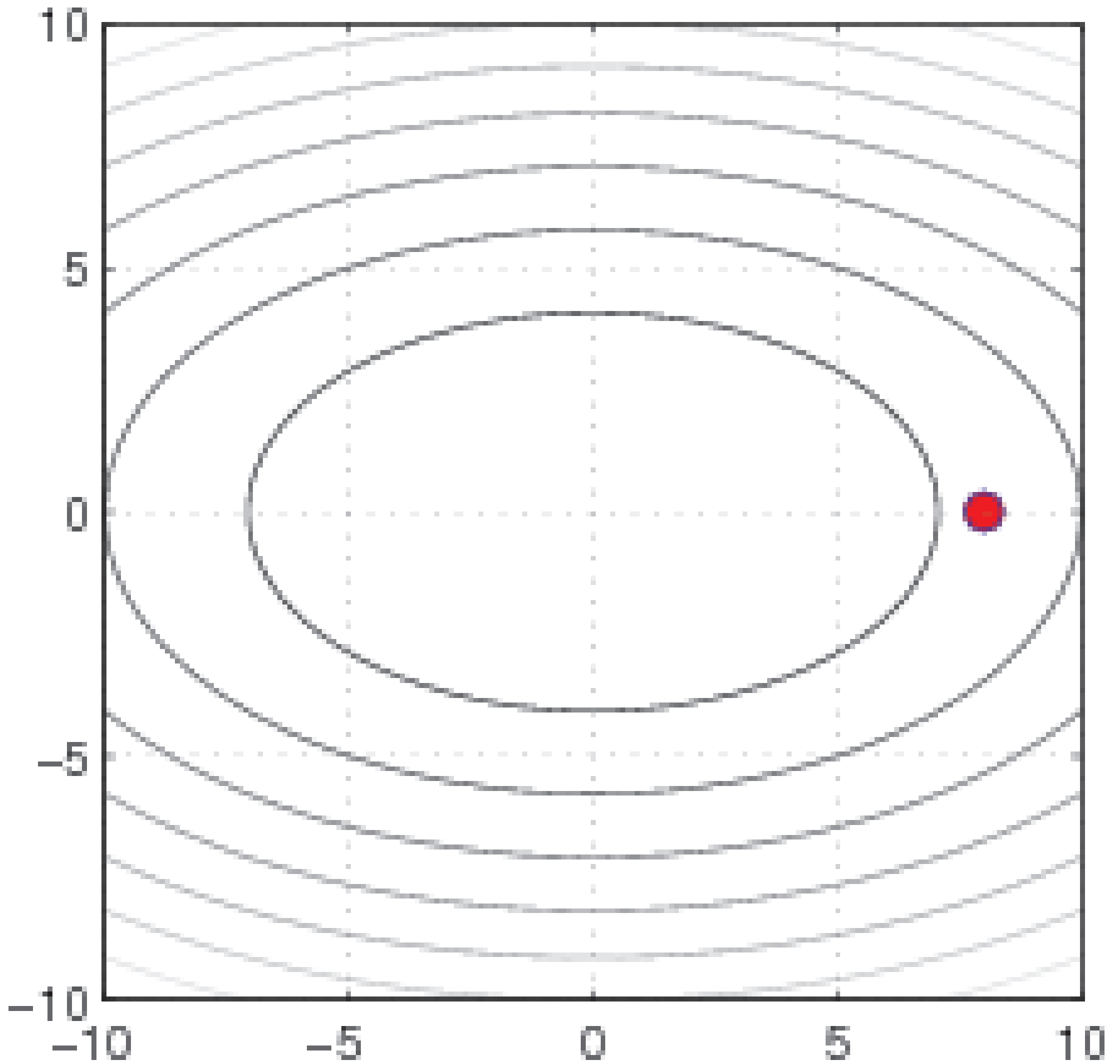
Note that, in the following contour plots, the coordinates of $x^{(0)}$ are indicated in the problem statement. The precise position may be important, so don't just eyeball it!

2.1 $f(x) = x_1^2 + x_2^2$ with $x^{(0)} = (8, 4)$



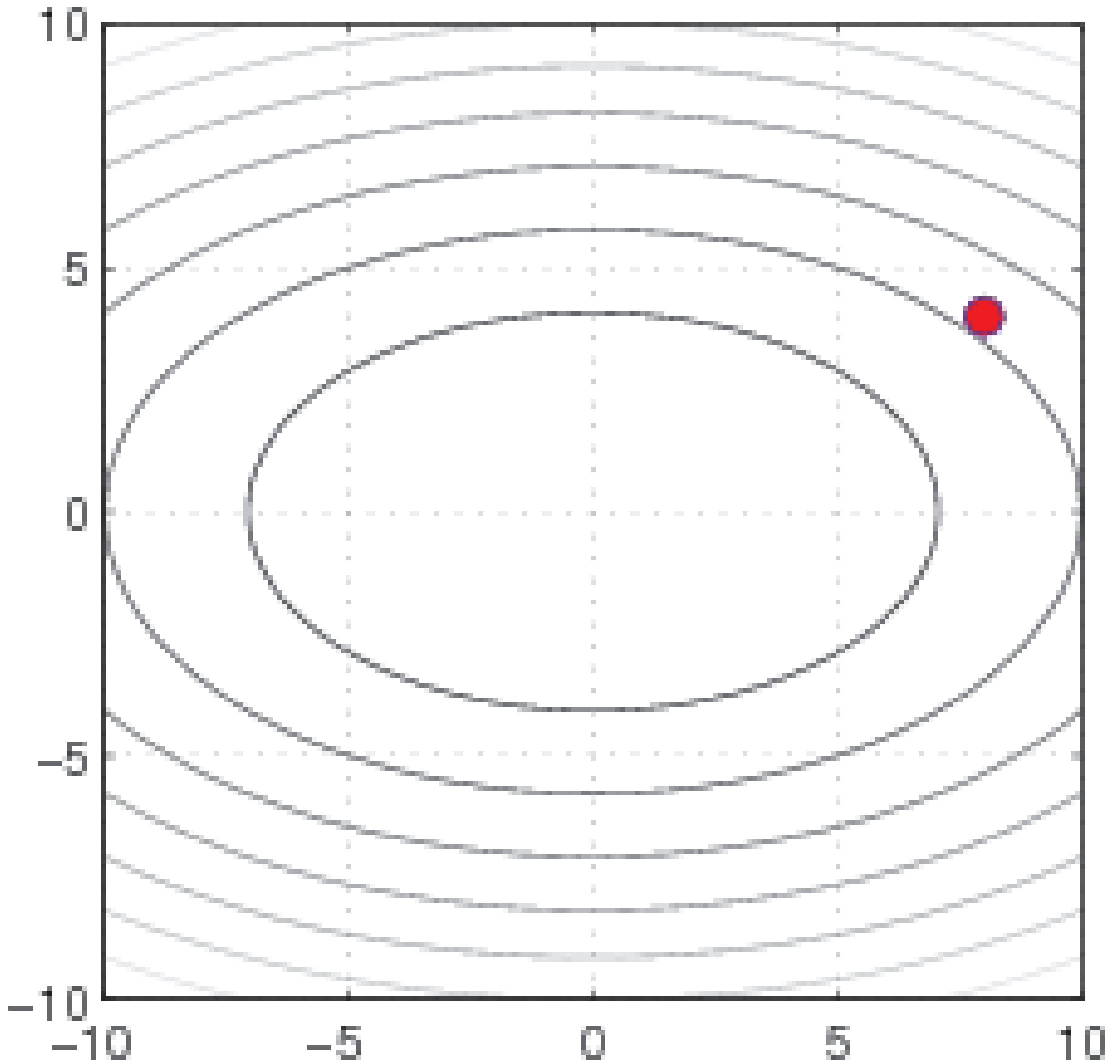
1. Gradient descent with exact line search $\quad x^{(1)} = \quad x^{(2)} = \quad x^{(3)} =$
2. Gradient descent with backtracking line search $\quad x^{(1)} = \quad x^{(2)} = \quad x^{(3)} =$
3. Newton's method with exact line search $\quad x^{(1)} = \quad x^{(2)} = \quad x^{(3)} =$
4. Steepest descent with respect to $\|\cdot\|_1$, with exact line search $\quad x^{(1)} = \quad x^{(2)} = \quad x^{(3)} =$

2.2 $f(x) = x_1^2 + 3x_2^2$ with $x^{(0)} = (0, 8)$



1. Gradient descent with exact line search $\quad x^{(1)} = \quad x^{(2)} = \quad x^{(3)} =$
2. Gradient descent with backtracking line search $\quad x^{(1)} = \quad x^{(2)} = \quad x^{(3)} =$
3. Newton's method with exact line search $\quad x^{(1)} = \quad x^{(2)} = \quad x^{(3)} =$
4. Conjugate gradient descent with exact line search $\quad x^{(1)} = \quad x^{(2)} = \quad x^{(3)} =$
5. BFGS with exact line search $\quad x^{(1)} = \quad x^{(2)} = \quad x^{(3)} =$

2.3 $f(x) = x_1^2 + 3x_2^2$ with $x^{(0)} = (8, 4)$



1. Gradient descent with exact line search $x^{(1)} =$ $x^{(2)} =$ $x^{(3)} =$
2. Newton's method with exact line search $x^{(1)} =$ $x^{(2)} =$ $x^{(3)} =$
3. Conjugate gradient descent with exact line search $x^{(1)} =$ $x^{(2)} =$ $x^{(3)} =$
4. BFGS with exact line search $x^{(1)} =$ $x^{(2)} =$ $x^{(3)} =$
5. Steepest descent with respect to $\|\cdot\|_1$, with exact line search $x^{(1)} =$ $x^{(2)} =$ $x^{(3)} =$

3

This question concerns optimization of the function

$$f(x) = \sum_{i=1}^{m} e^{a_i^T x_i + b_i}$$

for $x \in \mathbb{R}^n$.

3.1

Find the runtime complexity (as a function of m and n) to evaluate each of the following:

1. $f(x)$ Runtime: $O($ $)$

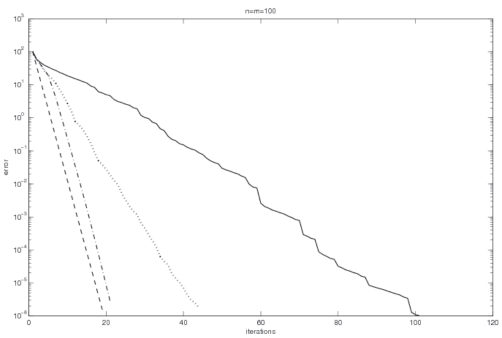
2. $\nabla f(x)$ Runtime: $O($ $)$

3. $\nabla^2 f(x)$ Runtime: $O($ $)$

3.2

In each of the following pairs of plots, identify the line style corresponding to each of the listed methods. The line search used for all algorithms was that of programming assignment 4 (which gives a step length satisfying the strong Wolfe conditions).

3.2.1



Plots of log-error versus number of iterations, and time, for each of the four algorithms. In both plots, the dimension is $n = 100$, and the number of a_i, b_i pairs is $m = 100$

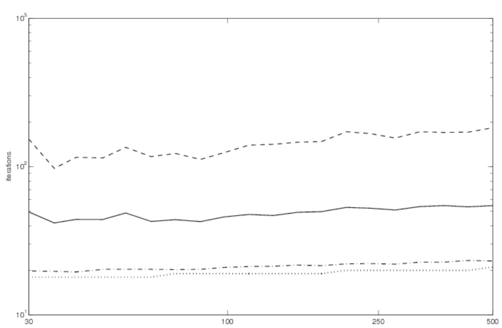
Dot — Gradient descent: ☐ Conjugate gradient: ☐ BFGS: ☐ Newton's method: ☐

Dash — Gradient descent: ☐ Conjugate gradient: ☐ BFGS: ☐ Newton's method: ☐

Dot-dash — Gradient descent: ☐ Conjugate gradient: ☐ BFGS: ☐ Newton's method: ☐

Solid — Gradient descent: ☐ Conjugate gradient: ☐ BFGS: ☐ Newton's method: ☐

3.2.2



Plots of the log-number-of-iterations and log-time until each algorithm achieves an error of 10^{-6}, versus $\ln n = \ln m$, where n is the problem dimension and m is the number of a_i, b_i pairs

Dot — Gradient descent: ☐ Conjugate gradient: ☐ BFGS: ☐ Newton's method: ☐

Dash — Gradient descent: ☐ Conjugate gradient: ☐ BFGS: ☐ Newton's method: ☐

Dot-dash — Gradient descent: ☐ Conjugate gradient: ☐ BFGS: ☐ Newton's method: ☐

Solid — Gradient descent: ☐ Conjugate gradient: ☐ BFGS: ☐ Newton's method: ☐

4

For each of a Linear Program, a Quadratic Program and a general convex program, indicate (by checking the appropriate box) whether each of the following situations is possible:

1. x in primal-feasible with $f(x) = 7$; λ, ν is dual-feasible with $g(\lambda, \nu) = 9$ LP: ☐ QP: ☐ Convex: ☐
2. x in primal-feasible with $f(x) = 7$; λ, ν is dual-feasible with $g(\lambda, \nu) = 5$ LP: ☐ QP: ☐ Convex: ☐
3. x in primal-optimal with $f(x) = 7$; λ, ν is dual-optimal with $g(\lambda, \nu) = 5$ LP: ☐ QP: ☐ Convex: ☐
4. The primal problem is infeasible; λ, ν is dual-feasible with $g(\lambda, \nu) = 5$ LP: ☐ QP: ☐ Convex: ☐
5. The primal problem is infeasible; λ, ν is dual-optimal with $g(\lambda, \nu) = 5$ LP: ☐ QP: ☐ Convex: ☐
6. The primal problem is infeasible; the dual problem is unbounded LP: ☐ QP: ☐ Convex: ☐

5

Consider the problem:

$$\begin{aligned} \text{minimize} &: x \\ \text{subject to} &: x^2 \leq 1 \end{aligned}$$

For parts 4-8, use the axis system on the next page.

1. What is the optimal value? $p^* =$

2. What is the Lagrangian?

3. What is the dual problem? (*Hint*: you should get a QP)

4. Recall the geometric view of duality. Plot the set:

$$\mathcal{G} = \{(u,t) \mid \exists x \left(f_0(x) = t \wedge f_1(x) = u\right)\}$$

   where $f_0(x)$ is the objective and $f_1(x) \leq 0$ is the constraint

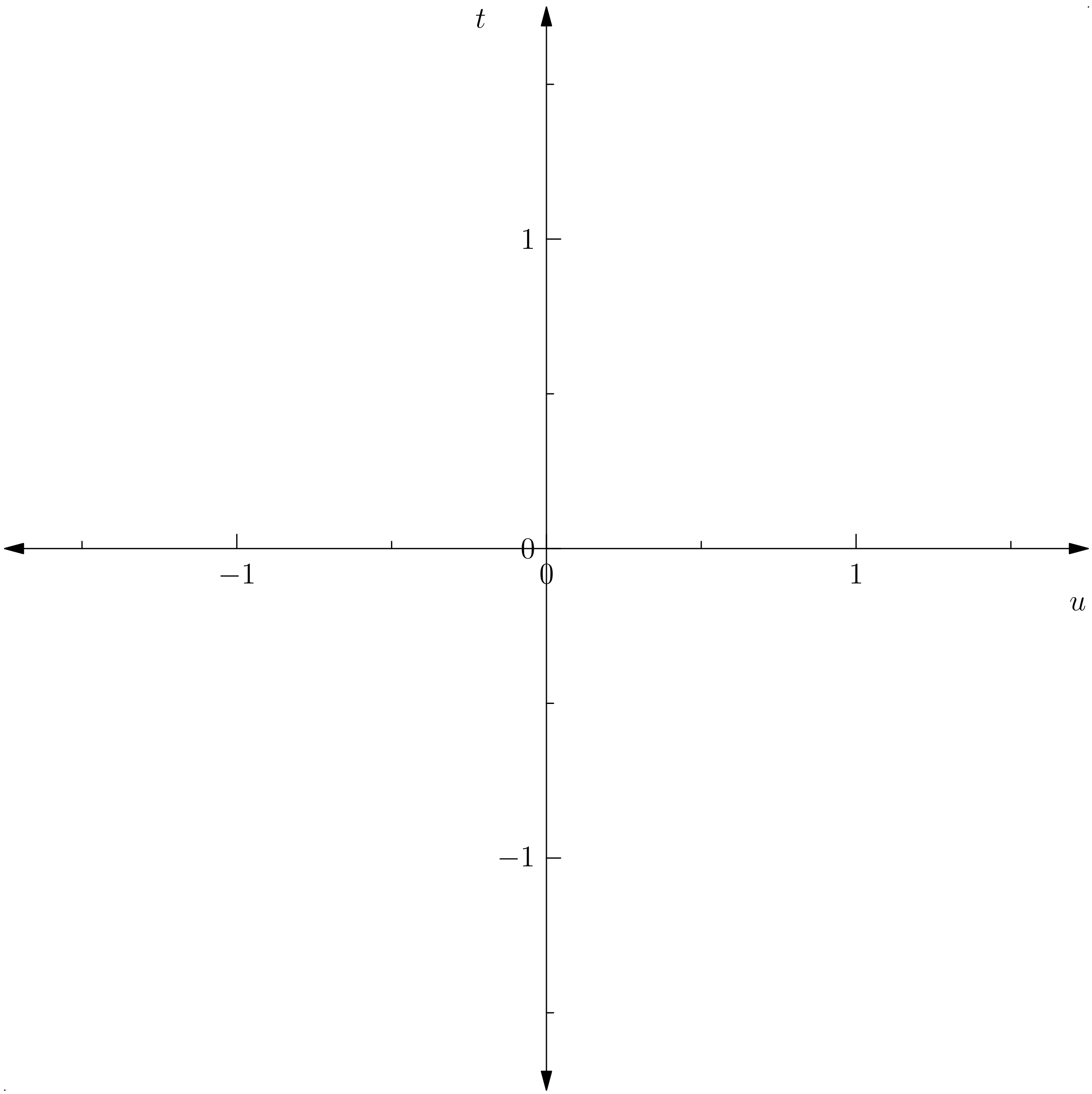
5. On the same figure, also show the set:

$$\mathcal{A} = \{(u,t) \mid \exists x \left(f_0(x) \leq t \wedge f_1(x) \leq u\right)\}$$

6. Draw the supporting line of $\mathcal{A}$ corresponding to the dual feasible solution $\lambda = 1$

7. What are the coordinates of the point where this line crosses the t axis? $(u,t) = (\quad , \quad)$
   Label this point on your plot

8. What are the coordinates of the point where this line touches A? $(u,t) = (\quad , \quad)$
   Label this point on your plot


t
1
0
−1
−1
0
1
u