Practice Test

Convex Optimization, 10-725

Name:

Andrew ID:

Each question is either in true/false format, or multiple choice. For multiple choice, just choose the single best option. In each case, make sure to fill in the box according to the answer you choose (true or false, or the multiple choice option) completely. All questions are worth 1 point.

1. The only functions that are both convex and concave are affine functions.
   ☐ True
   ☐ False

2. The least squares loss $f(\beta) = \|y - X\beta\|_2^2$ is always strongly convex.
   ☐ True
   ☐ False

3. Suppose we are minimizing a generic twice differentiable function f over $\mathbb{R}^n$ (its Hessian is assumed to be dense at each $x \in \mathbb{R}^n$). Both DFP and BFGS improve over the cost of Newton's method, per iteration, by a factor of:
   ☐ a. $O(n)$;
   ☐ b. $O(n^2)$;
   ☐ c. $O(\sqrt{n})$;
   ☐ d. they have the same cost per iteration.

4. The augmented Lagrangian parameter ρ in ADMM is really a theoretical formality, and the choice of ρ does not have any practical implications.
   ☐ True
   ☐ False

5. For a convex optimization problem with criterion f, a necessary and sufficient condition for a feasible point $x^\star$ to be optimal is that $g^T(x - x^\star) \geq 0$ for some $g \in \partial f(x^\star)$, and all feasible x.
   ☐ True
   ☐ False

6. In the barrier method, at any point $x(t)$ along the central path we can always construct points $u(t), v(t)$ that are feasible for the dual of the original optimization problem.
   ☐ True
   ☐ False

7. In any optimization problem, the criterion $f(x)$ can always be replaced by $e^{f(x)}$ without changing the solution.
   ☐ True
   ☐ False

8. In general, the minimizations over the two blocks of primal variables in ADMM can be done in parallel.
   ☐ True
   ☐ False

9. A Cholesky decomposition is well-defined for:
   ☐ a. any nonsingular matrix;
   ☐ b. any positive semidefinite matrix;
   ☐ c. any positive definite matrix;
   ☐ d. any matrix.

10. For the problem

$$\min_{x,y}\ f(x) + g(y),$$

where f, g are convex and f is strictly convex, the x-component of the solution must be unique. That is, if (x, y) and $(\tilde{x}, \tilde{y})$ are both solutions, then we must have $x = \tilde{x}$.
☐ True
☐ False

11. For a generic convex function, coordinatewise optimality implies global optimality.
☐ True
☐ False

12. A conic hull of points is a convex set.
☐ True
☐ False

13. For a strongly convex function f on $\mathbb{R}$, we cannot have $f(x) \to -\infty$ as $x \to \infty$.
☐ True
☐ False

14. For a strictly convex and twice differentiable function f, which is the following is true?
☐ a. $\text{dom}(f)$ is convex;
☐ b. $-f$ is strictly concave;
☐ c. $\nabla^2 f(x) \succ 0$ at all x;
☐ d. both (a) and (b);
☐ e. all of (a), (b), (c).

15. The SVM problem satifies strong duality.
☐ True
☐ False

16. An advantage of the barrier method over the primal-dual interior point method is that:
☐ a. the barrier method generally converges faster;
☐ b. the barrier method yields dual feasible points throughout;
☐ c. the barrier method yields primal feasible points throughout;
☐ d. the barrier method maintains complementary slackness throughout.

17. The polyhedron $\{x : Ax \leq b\}$ is always convex, regardless of A.
☐ True
☐ False

18. Stochastic gradient descent generally converges more rapidly to a high accuracy solution than does its nonstochastic counterpart.
☐ True
☐ False

19. A nonconvex optimization problem always has multiple local minima.
☐ True
☐ False

20. Stochastic gradient descent on a strongly convex function with a Lipschitz gradient, converges (with suitable step sizes) at the rate:
☐ a. $O(1/\epsilon)$;
☐ b. $O(1/\sqrt{\epsilon})$;
☐ c. $O(1/\epsilon^2)$;
☐ d. $O(\log(1/\epsilon))$.

21. A convex function must be defined on all of $\mathbb{R}^n$.
☐ True
☐ False

22. If f has convex sublevel sets, $\{x : f(x) \leq \alpha\}$, for all $\alpha \in \mathbb{R}$, then f is convex.
☐ True
☐ False

23. The surrogate duality gap in the primal-dual interior-point method is a bonafide duality gap when both the primal and dual residuals are zero.
☐ True
☐ False

24. The SR1 update ensures that the approximated Hessian remains positive definite.
☐ True
☐ False

25. In the barrier method, if there are m inequality constraints, then the suboptimality gap at a point $x(t)$ on the central path is bounded above by:
☐ a. m/t;
☐ b. t/m;
☐ c. $\sqrt{t}/m$;
☐ d. there is no general closed-form inolving t, m.

26. The conjugate of $f(x) = x^2$ is:
☐ a. $f^*(y) = y^2$;
☐ b. $f^*(y) = y^2/2$;
☐ c. $f^*(y) = 2y^2$;
☐ d. $f^*(y) = y^2/4$.

27. For minimizing a convex, differentiable function f, whose gradient is Lipschitz, gradient descent achieves the optimal rate among first-order methods.
☐ True
☐ False

28. For minimizing the least squares loss $\|y - X\beta\|_2^2$, the coordinate descent updates are equivalent to the Gauss-Seidel updates for solving the linear system $X^T X\beta = X^T y$.
☐ True
☐ False

29. If f_1, f_2 are convex and differentiable, then subgradients of the function $f = \max\{f_1, f_2\}$ at a point x such that $f_1(x) = f_2(x)$ are:
☐ a. only $\nabla f_1(x)$ and $\nabla_2 f(x)$;
☐ b. all convex combinations of $\nabla f_1(x)$ and $\nabla f_2(x)$;
☐ c. all linear combinations of $\nabla f_1(x)$ and $\nabla f_2(x)$;
☐ d. none of the above.

30. Projection onto the feasible set in a linear program can always be done in closed-form.
☐ True
☐ False

31. The Frank-Wolfe method is affine invariant.
☐ True
☐ False

32. Both DFP and BFGS converge at the same (local) rate as Newton's method, under the same set of assumptions.
☐ True
☐ False

33. When comparing QR and Cholesky decompositions to solve a least squares problem, generally speaking, it holds that:
☐ a. QR is cheaper, Cholesky is more stable;

☐ b. QR is cheaper and more stable;
☐ c. QR is more stable, Cholesky is cheaper;
☐ d. none of the above.

34. In coordinate descent, after minimizing over coordinate i, either the new value or the old value for this coordinate can be used for the minimization over coordinate $i + 1$; either choice will result in a convergent algorithm.
☐ True
☐ False

35. Newton's method is affine invariant.
☐ True
☐ False