quiz

convex optimization

sp1

A function f is continuous differentiable, that is $f \in C^1$, if

Vælg en eller flere:

- [ ] f has continuous first order partial derivatives
- [ ] f is differentiable
- [x] f is differentiable and its derivative is continuous
- [ ] f has first order partial derivatives
- [ ] f is continuous and its partial derivatives exists

sp2

The gradient ∇f of a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is

Vælg en eller flere:

- [x] a scalar if $n = 1$ ✓
- [ ] a vector with coordinates equal to the second order partial derivatives of f
- [ ] a scalar
- [ ] a square $n \times n$ matrix
- [x] a vector with coordinates equal to the first order partial derivatives of f ✓
- [ ] the derivative of f
- [x] a n-vector ✓

sp3

Let $x \in \mathbb{R}^n$. The Hessian $H(x)$ of a C^2 function $f : \mathbb{R}^n \to \mathbb{R}$ is

Vælg en eller flere:

- [x] a square $n \times n$ matrix ✔
- [ ] a scalar if $n = 1$
- [ ] a vector with coordinates equal to the second order partial derivatives of f
- [ ] a n-vector
- [ ] a scalar
- [x] the Jacobian matrix of ∇f ✔
- [ ] a vector with coordinates equal to the first order partial derivatives of f
- [ ] the Jacobian matrix of f
- [x] a symmetric matrix ✔
- [x] the second derivative of f ✔
- [ ] the derivative of f

sp4

A linear approximation of the function f involve

Vælg en eller flere:

- [ ] the Hessian of f
- [x] the gradinet of f ✔

sp5

A quadratic approximation of the function f involve

Vælg en eller flere:

- [x] the gradinet of f
- [x] the Hessian of f ✔

Here's why:

The **quadratic (second-order Taylor) approximation** of a scalar function $f : \mathbb{R}^n \to \mathbb{R}$ around a point $\mathbf{x}_0$ is:

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T H_f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$$

Where:

- $\nabla f(\mathbf{x}_0)$: the **gradient** of f at $\mathbf{x}_0$ — this gives the **linear (first-order)** behavior.
- $H_f(\mathbf{x}_0)$: the **Hessian** of f at $\mathbf{x}_0$ — this gives the **curvature (second-order)** information.

So:

- □ **Gradient** is needed — tells us the slope (first-order).
- □ **Hessian** is needed — tells us how the slope changes (second-order).

In summary:

To build a **quadratic approximation**, you need:

- The **function value** at the point,
- The **gradient** (first derivatives),
- The **Hessian** (second derivatives).

sp6

A point $x' \in \mathbb{R}^n$ is a weak local minimizer of $f : \mathbb{R}^n \to \mathbb{R}$ if

Vælg en eller flere:

- [ ] $f(x) \geq f(x')$ for all x
- [ ] $f(x) \leq f(x')$ for all x sufficiently close to x'
- [ ] $f(x) \leq f(x')$ for all x
- [x] $f(x) \geq f(x')$ for all x sufficiently close to x'

sp7

A point $x' \in \mathbb{R}^n$ is a weak global minimizer of $f : \mathbb{R}^n \to \mathbb{R}$ if

Vælg en eller flere:

- [x] $f(x) \geq f(x')$ for all x ✓
- [ ] $f(x) \leq f(x')$ for all x
- [ ] $f(x) \leq f(x')$ for all x sufficiently close to x'
- [ ] $f(x) \geq f(x')$ for all x sufficiently close to x'

sp8

If x' is a local minimizer of $f : \mathbb{R}^n \to \mathbb{R}$, with $f \in C^1$, then

Vælg en eller flere:

- [x] the gradient of f at x' is zero ✓
- [ ] the gradient of f at x' is smaller than or equal to zero
- [ ] the gradient of f is zero
- [ ] the gradient of f at x' is bigger than or equal to zero

� **Intuition:**

- The gradient ∇f(x′) points in the direction of steepest increase.
- If you're at a **local minimum**, there's **no direction** to go that would decrease the function further (at least locally).
- Therefore, the slope in **every direction** is zero — and this is exactly what the gradient being zero means.

sp9

If x' is a local minimizer of $f : \mathbb{R}^n \to \mathbb{R}$, with $f \in C^2$, then

Vælg en eller flere:

- [ ] the gradient of f at x' is zero or the hessian of f at x' is positive semi definite
- [x] the gradient of f at x' is zero and the hessian of f at x' is positive semi definite
- [ ] the gradient of f at x' is zero or the hessian of f at x' is positive definite
- [ ] the gradient of f at x' is zero and the hessian of f at x' is positive definite

sp10

A set $S \subseteq \mathbb{R}^n$ is convex if

Vælg en eller flere:

- [x] $S = \{x \in \mathbb{R}^n \mid ||x||_1 \leq 5\}$
- [ ] a curve between any two points in S is contained in S
- [ ] $S = \{x \in \mathbb{R}^n \mid ||x||_2 = 1\}$ is a circle
- [x] the straight line between any two points in S is contained in S
- [ ] $n = 1$ and S is the union of two disjoint intervals
- [x] $S = \{x \in \mathbb{R}^n \mid ||x||_2 \leq 1\}$ is a disk

✅ 1. $S = \{x \in \mathbb{R}^n \mid \|x\|_1 \leq 5\}$

- Yes — this set is **convex.**
- It's the **unit ball** in ℓ^1-norm (scaled by 5).
- In **2D**, this looks like a **diamond shape** (not a box), with corners at $(5, 0), (0, 5), (-5, 0), (0, -5)$.
- In **higher dimensions**, it's a **cross-polytope**, not a box.
- ⚠️ **So, it's *not* a box**, but **it is convex.**

sp11

A function $f : R \to \mathbb{R}$, with $R \subseteq \mathbb{R}^n$ and $f \in C^2$, is convex if

Vælg en eller flere:

- [x] f is a quadratic function
- [ ] R is convex and the hessian $H(x)$ of f is positive semi definite for some $x \in R$
- [x] R is convex and the hessian $H(x)$ of f is positive semi definite for all $x \in R$
- [ ] the hessian $H(x)$ of f is positive semi definite for all $x \in R$
- [ ] R is convex
- [ ] f is a linear function

✅ Correct answer 2:

"f is a linear function"

- Linear functions are of the form $f(x) = a^T x + b$.
- Linear functions are **both convex and concave**, because:

$$f(\lambda x + (1-\lambda)y) = \lambda f(x) + (1-\lambda)f(y)$$

- This is **stronger** than the convexity inequality, so linear functions are always convex.

sp12

Let $R \subseteq \mathbb{R}^n$. The optimization problem $\min_{x \in R} f(x)$ is convex if

Vælg en eller flere:

- [x] R is convex and $f : R \to \mathbb{R}$ is convex ✔
- [ ] $f : S \to \mathbb{R}$ is convex for some (convex) set S with $R \subset S$
- [ ] R is convex
- [ ] R is convex and $f : S \to \mathbb{R}$ is convex for some (convex) set S with $S \subset R$
- [ ] $f : R \to \mathbb{R}$ is convex
- [x] R is convex and $f : S \to \mathbb{R}$ is convex for some (convex) set S with $R \subset S$ ✔

✅ Correct Answer 2:

"$f : R \to \mathbb{R}$ **is convex**"

- This is **necessary**, but not **sufficient** on its own.
- However, it is still considered **part of what makes a problem convex**.
- If this option was presented alone, it might not be enough — but in the context of multiple correct answers, it's valid as a **partial requirement**.

✔ **This is correct, assuming R is convex too.**

✅ Correct Answer 3:

"R is convex and $f : S \to \mathbb{R}$ is convex for some (convex) set S with $R \subset S$"

- This is also valid. If f is convex on a **larger convex set** S that contains R, and R is convex, then:
  - f is convex **on** R as well.
- Convexity is preserved when you **restrict a convex function to a convex subset** of its domain.

✔ So this also guarantees a convex optimization problem.

🔤 In plain language:

$f : R \to \mathbb{R}$ means:

- There is a **function** f,
- It takes **inputs from the set** R,
- And it **outputs real numbers** (i.e., values in $\mathbb{R}$).

sp13

Mark the statements which are true:

Vælg en eller flere:

- [x] The software tool CVX is a numerical solver for optimization problems ✖
- [ ] Any optimization problem can be approximated arbitrary well with a convex optimization problem
- [x] The solution set of a convex optimization problem is convex ✔
- [ ] Local minimizers are global minimizers in general
- [x] For convex optimization problems local minimizers are global minimizers ✔
- [ ] Optimization problems are in general numerically solvable
- [x] The software tool CVX is a numerical solver for convex optimization problems ✔
- [x] Convex optimization problems are in general numerically solvable ✔

gradient methods

sp1

The steepest-descent method can be applied to

Vælg en eller flere:

- [x] least squares problems ✔
- [ ] constrained optimization problems
- [x] unconstrained convex optimization problems ✔
- [ ] unconstrained optimization problems ✓

✅ **1. Unconstrained optimization problems**

- **Steepest descent** is a method for **minimizing a function** without any constraints on the variables.
- It uses the **negative gradient direction** (steepest decrease) to iteratively update the variable x.
- So, **any optimization problem without constraints** is a valid candidate.

📌 Correct

✅ 2. Unconstrained convex optimization problems

- If the function is **convex** and **differentiable**, the **steepest descent method** is **guaranteed to converge** to a **global minimum**, assuming proper step sizes.
- Convexity makes the method **more reliable and efficient**.

📌 **Correct and even better than the general case**

sp2

The steepest-descent method is a

Vælg en eller flere:

- [x] first order method
- [ ] third order method
- [ ] fourth order method
- [ ] second order method

sp3

Let $f \in C^2$ be a cost function, H its hessian, and $\{x_k\}$ denote the sequence generated by the steepest-descent algorithm. For x_k close to a minimizer x^* the algorithm converge fast to x^* if

Vælg en eller flere:

- [ ] the eigenvalues of $H(x_k)$ are almost the same
- [ ] the eigenvalues of $H(x_k)$ are small
- [ ] all eigenvalues of $H(x_k)$ are real
- [x] the largest and smallest eigenvalue of $H(x_k)$ are almost the same

(condition Number)

sp4

The Steepest-Descent Algorithm (SDA) and the Newton-Raphson Algorithm (NRA) can be used in combination by

Vælg en eller flere:

- [ ] using the NRA fare from the minimizer and the SDA close to the minimizer
- [x] using the SDA fare from the minimizer and the NRA close to the minimizer
- [ ] iterating between the SDA and the NRA

sp5

The Newton-Raphson method is a

Vælg en eller flere:

- [x] second order method
- [ ] fourth order method
- [ ] third order method
- [ ] first order method

sp6

When attempting to solve $0 = \nabla f(x)$ for x, one can apply

Vælg en eller flere:

- [ ] a first order method
- [x] the Gauss-Newton method
- [ ] the steepest-descent method
- [ ] the Newton-Raphson method

🔍 What does $0 = \nabla f(x)$ mean?

You're looking for a point x where the **gradient of** f is zero — in other words, a **stationary point**, which could be a minimum, maximum, or saddle point.

✅ Final Summary:

Reason Gauss-Newton is good	Explanation
Designed for least squares	Solves problems like $\min \|r(x)\|^2$
Solves $\nabla f(x) = 0$	Because $\nabla f(x) = J^T r$
Uses Jacobians only	No need for second derivatives (Hessian)
Faster than Newton	Approximates Hessian with $J^T J$
Converges well	Especially when residuals are small

constrained optimization

PLEASE WRITE WHY?!?!?!?!?!?

sp1

Consider the constrained optimization problem $\min f(x) \quad s.t \quad \mathbf{a}(x) = (a_1(x), \dots, a_p(x)) = 0$, and let $J(x) = J_{\mathbf{a}}(x)$ denote the Jacobian of $\mathbf{a}$ at x. The point x' is regular if

Vælg en eller flere:

- $\mathbf{a}(x') = 0$ and the gradients $\nabla a_1(x'), \dots, \nabla a_p(x')$ are linear independent
- $a_i(x') = 0$ for $i = 1, \dots, p$ and $J(x)$ has full row rank
- $Rank(A) = p$ *in the case where* $\mathbf{a}(x) = Ax - b,\ A \in \mathbb{R}^{p \times n}$
- $\mathbf{a}(x') = 0$ and some of the gradients $\nabla a_1(x'), \dots, \nabla a_p(x')$ are linear independent
- $\mathbf{a}(x') = 0$ and $J(x')$ has full row rank
- $J(x')$ has full row rank
- the gradients $\nabla a_1(x'), \dots, \nabla a_p(x')$ are linear independent
- $Ax' = b$ and $Rank(A) = p$, *in the case where* $\mathbf{a}(x) = Ax - b,\ A \in \mathbb{R}^{p \times n}$
- $a_i(x') = 0$ for some i and $J(x)$ has full row rank

sp2

Consider the constrained optimization problem $\min f(x) \quad s.t \quad \mathbf{a}(x) = (a_1(x), \dots, a_p(x)) = 0,\ \mathbf{c}(x) = (c_1(x), \dots, c_q(x)) \geq_e 0$, and let $J_{\mathbf{a}}(x),\ J_{\mathbf{c}}(x)$ denote the Jacobian of $\mathbf{a},\ \mathbf{c}$ at x. The point x' is regular if

Vælg en eller flere:

- $\mathbf{a}(x') = 0,\ \mathbf{c}(x') \geq_e 0$ and the gradients $\nabla a_1(x'), \dots, \nabla a_p(x'),\ \nabla c_{j_1}(x'), \dots, \nabla c_{j_k}(x')$ are linearly independent, with $c_{j_l}(x') = 0$ for $l = 1, \dots k$
- $\mathbf{a}(x') = 0,\ \mathbf{c}(x') \geq_e 0$ and the gradients $\nabla a_1(x'), \dots, \nabla a_p(x'),\ \nabla c_{j_1}(x'), \dots, \nabla c_{j_k}(x')$ are linearly independent, with $c_{j_l}(x') > 0$ for $l = 1, \dots k$
- $\mathbf{a}(x') = 0,\ \mathbf{c}(x') \geq_e 0$ and the gradients $\nabla a_1(x'), \dots, \nabla a_p(x')$ are linearly independent
- $\mathbf{a}(x') = 0,\ \mathbf{c}(x') \geq_e 0$ and the gradients $\nabla a_1(x'), \dots, \nabla a_p(x'),\ \nabla c_{j_1}(x'), \dots, \nabla c_{j_k}(x')$ are linearly independent, with $c_{j_l}(x') \geq 0$ for $l = 1, \dots k$
- $\mathbf{a}(x') = 0,\ \mathbf{c}(x') \geq_e 0$ and $J_{\mathbf{a}}(x')$ has full row rank
- $a_i(x') = 0$ for $i = 1, \dots, p$, $c_i(x') \geq 0$ for $i = 1, \dots, q$ and $J_{\mathbf{a}}(x)$ has full row rank

sp3

Consider the constrained optimization problem $\min f(x) \quad s.t \quad \mathbf{a}(x) = (a_1(x), \dots, a_p(x)) = 0$, and let $J(x') = J_{\mathbf{a}}(x')$ denote the Jacobian of $\mathbf{a}$ at x'. Let the point x' be a regular minimizer (with $\nabla f(x') \neq 0$)

Vælg en eller flere:

- then $\nabla f(x')$ is in the range of $J(x')$
- then $\nabla f(x')$ is orthogonal to the gradients $\nabla a_1(x'), \dots, \nabla a_p(x')$
- then $\nabla f(x')$ is in the range of $J(x')^\top$
- then $\nabla f(x'),\ \nabla a_1(x'), \dots, \nabla a_p(x')$ are linearly independent
- then $\nabla f(x')$ can be expressed as a linear combination of the gradients $\nabla a_1(x'), \dots, \nabla a_p(x')$
- then $\nabla f(x')$ is in the image of $J(x')^\top$
- then $\nabla f(x')$ is in the image of $J(x')$

sp4

Consider the constrained optimization problem $\min f(x) \quad s.t \quad \mathbf{a}(x) = (a_1(x), \dots, a_p(x)) = 0,\ \mathbf{c}(x) = (c_1(x), \dots, c_q(x)) \geq_e 0$. Let the point x' be a regular minimizer (with $\nabla f(x') \neq 0$)

Vælg en eller flere:

- then $\nabla f(x')$ is a linear combination of $\nabla a_1(x'), \dots, \nabla a_p(x'),\ \nabla c_{j_1}(x'), \dots, \nabla c_{j_k}(x')$ with $c_{j_l}(x') = 0$ for $l = 1, \dots k$
- then $\nabla f(x')$ is a linear combination of $\nabla a_1(x'), \dots, \nabla a_p(x'),\ \nabla c_{j_1}(x'), \dots, \nabla c_q(x')$
- then $\nabla f(x')$ is a linear combination of $\nabla a_1(x'), \dots, \nabla a_p(x'),\ \nabla c_{j_1}(x'), \dots, \nabla c_{j_k}(x')$ with $c_{j_l}(x') > 0$ for $l = 1, \dots k$
- then $\nabla f(x')$ is a linear combination of $\nabla a_1(x'), \dots, \nabla a_p(x'),\ \nabla c_{j_1}(x'), \dots, \nabla c_{j_k}(x')$ with $c_{j_l}(x') \geq 0$ for $l = 1, \dots k$

sp5

Consider the constrained optimization problem $\min f(x) \quad s.t \quad \mathbf{a}(x) = (a_1(x), \ldots, a_p(x)) = 0,\ \mathbf{c}(x) = (c_1(x), \ldots, c_q(x)) \geq_e \mathbf{0}$. Let the point x' be a regular minimizer and $\lambda',\ \mu'$ be the corresponding multipliers

Vælg en eller flere:

- [ ] If $\mu_i > 0$ then $c_i(x') > 0$
- [ ] If $\mu_i > 0$ then $c_i(x') = 0$
- [ ] If $\mu_i > 0$ then $c_i(x') < 0$
- [x] If $c_i(x') > 0$ then $\mu_i = 0$
- [ ] If $\nabla f(x') = 0$ then $\lambda' \neq 0,\ \mu' = 0$
- [ ] If $c_i(x') > 0$ then $\mu_i \neq 0$
- [ ] If $\nabla f(x') = 0$ then $\lambda' = 0,\ \mu' > 0$
- [x] If $\nabla f(x') = 0$ then $\lambda' = 0,\ \mu' = 0$

sp6

Consider the constrained optimization problem $\min f(x) \quad s.t \quad \mathbf{a}(x) = (a_1(x), \ldots, a_p(x)) = 0,\ \mathbf{c}(x) = (c_1(x), \ldots, c_q(x)) \geq_e \mathbf{0}$. The KKT conditions are necessary and sufficient when

Vælg en eller flere:

- [x] the constrained optimization problem is convex
- [ ] $-f$ and $-\mathbf{c}$ are convex and $\mathbf{a}$ is affine linear
- [ ] f and $-\mathbf{c}$ are convex and $\mathbf{a}$ is affine linear
- [x] f, $\mathbf{a}$ and $-\mathbf{c}$ are convex
- [ ] f and $\mathbf{c}$ are convex and $\mathbf{a}$ is affine linear
- [ ] the constrained optimization problem is a linear programming problem