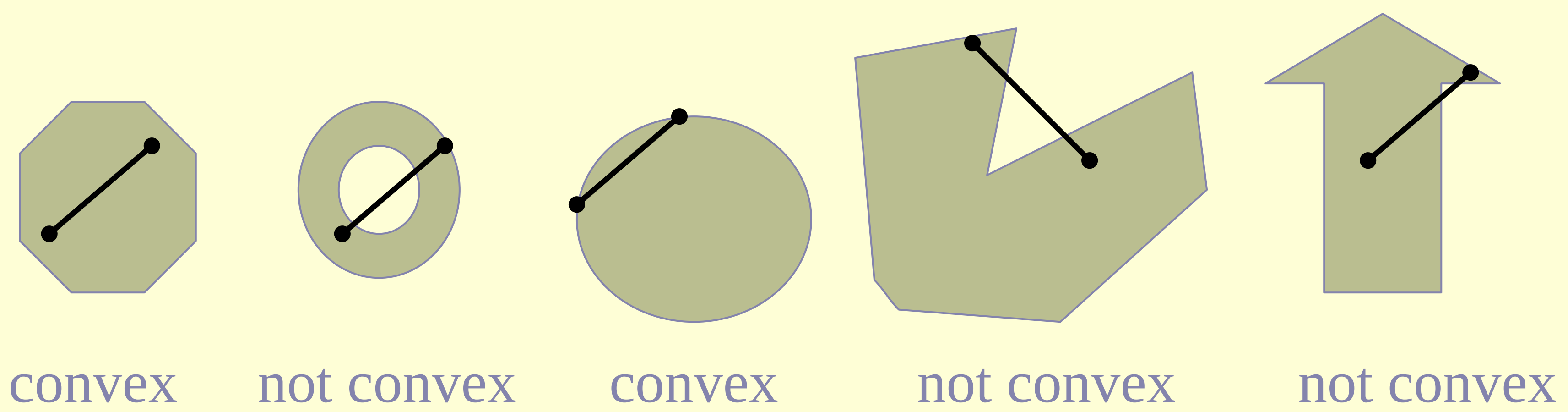
Computational Optimization

Convexity and
Unconstrained Optimization
1/29/08 and 2/1(revised)

Convex Sets

A set S is convex if the line segment joining any two points in the set is also in the set, i.e., for any $x,y \in S$,

$\lambda x+(1- \lambda)y \in S$ for all $0 \leq \lambda \leq 1$ }.



Proving Convexity

- Prove {x|Ax<=b} is convex.

Let x and y be elements of $C=\{x|Ax \le b\}$.

For any $\lambda \in (0,1)$,

$$A(\lambda x+(1-\lambda)y)=\lambda Ax+(1-\lambda)Ay$$

$$\le \lambda b+(1-\lambda)b=b$$

$$\Rightarrow \lambda x+(1-\lambda)y \in C$$

You Try

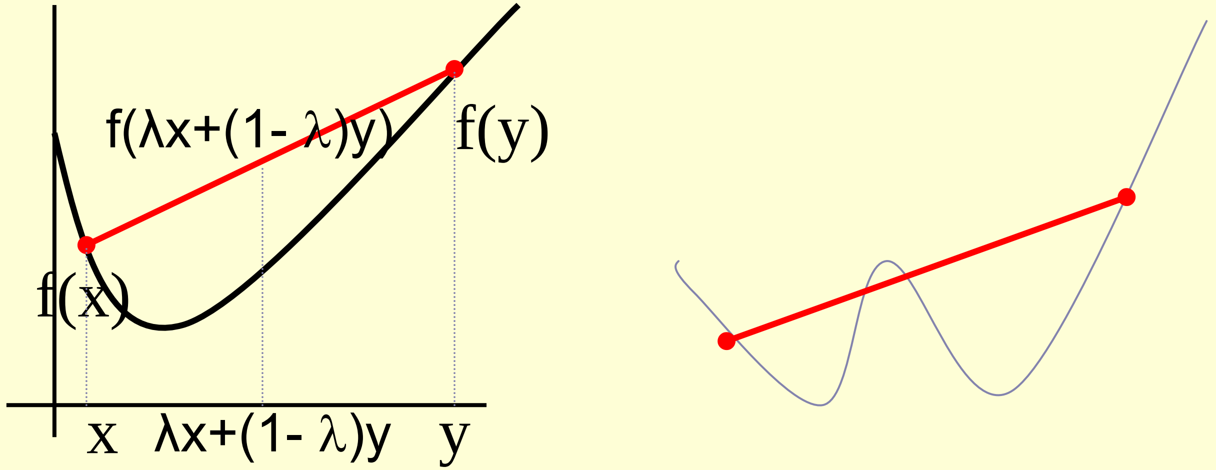
- Prove $D = \{x \mid \|x\| \leq 1\}$ is convex.

Convex Functions

A function f is (strictly) convex on a convex set S, if and only if for any $x, y \in S$,

$$f(\lambda x+(1-\lambda)y)(<) \le \lambda f(x)+(1-\lambda)f(y)$$

for all $0 \le \lambda \le 1$.



Proving Function Convex

- Linear functions

$$f(x) = w'x = \sum_{i=1}^{n} w_i x_i \quad where \;\; x \in R^n$$

$$For \; any \;\; x, y \in R^n \;\; \lambda \in (0,1)$$

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

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

$\square$

You Try

$$f(x_1, x_2) = x_1^{2} + 2x_2^2$$

Hint: x^2 is convex

Consider any two points x,y and $\lambda \in (0,1)$

$$\lambda x^2 + (1-\lambda) y^2 = \lambda^2 x^2 + (1-\lambda)^2 y^2 + (1-\lambda)\lambda x^2 + \lambda(1-\lambda) y^2$$

$$= \left(\lambda x + (1-\lambda) y\right)^2 - 2\lambda(1-\lambda)xy + (1-\lambda)\lambda x^2 + \lambda(1-\lambda) y^2$$

$$= \left(\lambda x + (1-\lambda) y\right)^2 + \lambda(1-\lambda)(x-y)^2$$

$$\geq \left(\lambda x + (1-\lambda) y\right)^2$$

First line uses $\lambda x^2 = \lambda^2 x^2 + (1-\lambda)\lambda x^2$ and similarly for $(1-\lambda) y^2$.

Second line completes the square of $\lambda^2 x^2 + (1-\lambda)^2 y^2$.

Third line observes the remaining terms are a square.

Fouth line follows since $\lambda(1-\lambda)(x-y)^2 \geq 0$.

Handy Facts

Let $g_1(x),\ldots,g_m(x)$ be convex functions

And a>0.

Then $f(x)=\sum_{i=1}^{m} g_i(x)$ is convex.

And

$h(x)=ag_1(x)$ is convex.

Convexity and Curvature

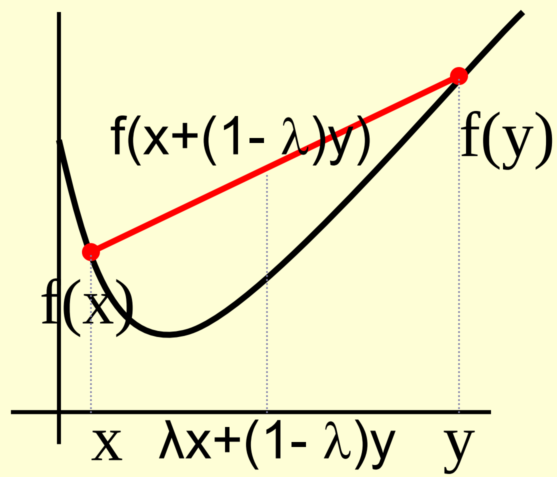
- Convex functions have positive curvature everywhere.
- Curvature can be measured by the second derivative or Hessian.
- Properties of the Hessian indicate if a function is convex or not.

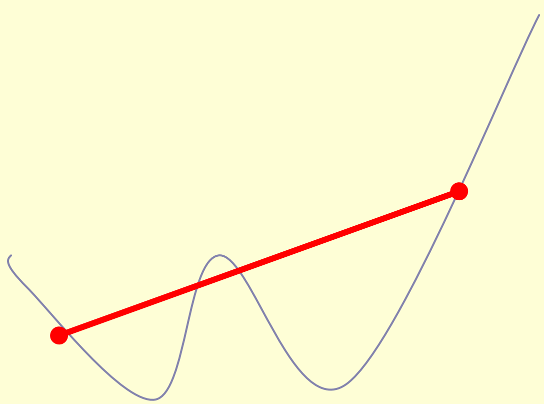
Convex Functions

A function f is (strictly) convex on a convex set S, if and only if for any $x, y \in S$,

$$f(\lambda x+(1-\lambda)y)(<) \leq \lambda f(x)+(1-\lambda)f(y)$$

for all $0 \leq \lambda \leq 1$.





Theorem

- Let f be twice continuously differentiable.

  f(x) is convex on S if and only if for all $x \in S$, the Hessian at x

$$\nabla^2 f(x)$$

  is positive semi-definite.

Definition

- The matrix H is positive semi-definite (p.s.d.) if and only if for any vector y

$$y'Hy \geq 0$$

- The matrix H is positive definite (p.d.) if and only if for any nonzero vector y

$$y'Hy > 0$$

- Similarly for negative (semi-) definite.

Theorem

- Let f be twice continuously differentiable.

  f(x) is *strictly* convex on S if and only if for all x∈X, the Hessian at x

$$\nabla^2 f(x)$$

is positive definite.

Checking Matrix H is p.s.d/p.d.

- Manually

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 4x_1^2 - x_2 x_1 - x_1 x_2 + 3x_2^2$$

$$= 4x_1^2 - 2x_1 x_2 + 3x_2^2$$

$$= (x_1 - x_2) \wedge 2 + 3x_1^2 + 2x_2^2 > 0 \ \forall [x_{1,} x_2] \neq 0$$

so matrix is positive definite

Useful facts

- The sum of convex functions is convex
- The composition of convex functions is convex.

via eigenvalues

- The eigenvalues of

$$\begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix} \text{are } 4.618 \text{ and } 2.382$$

so matrix is positive definite

Summary: using eigenvalues

- If all eigenvalues are positive, then matrix is positive definite, p.d.
- If all eigenvalues are nonnegative, then matrix is positive semi-definite, p.s.d
- If all eigenvalues are negative, then matrix is negative definite, n.d.
- If all eigenvalues are nonpositive, then matrix is negative semi-definite, n.s.d
- Otherwise the matrix is indefinite.

Try with Hessians

$$f(x_1, x_2) = x_1^{\,2} \;\; + 2x_2^2$$

$$\nabla f(x) = \begin{bmatrix} 2x_1 \\ 4x_2 \end{bmatrix}$$

$$\nabla^2 f(x) = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

$$[a\,b]\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = 2a^2 + 4b^2 > 0 \;\; for\ any\ [ab] \neq 0$$

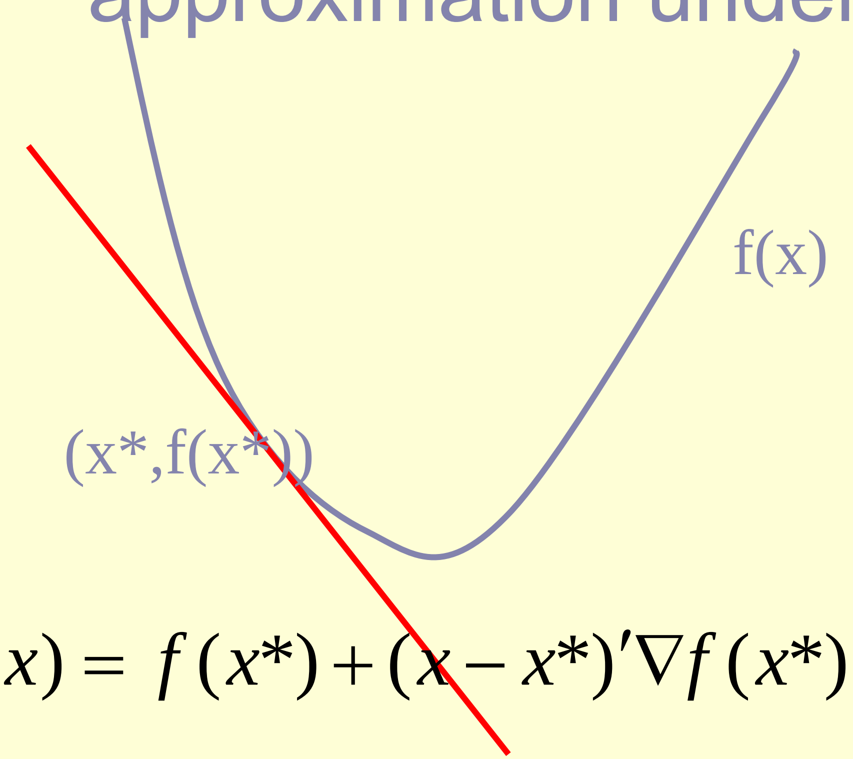
$$Strictly\ Convex$$

Check Hessian

- H = [2 0; 0 4]
- Eigs(H)   are 2 and 4
- So Hessian matrix is always p .d.
- So function is strictly convex

Differentiability and Convexity

- For convex function, linear approximation underestimates function



$$g(x) = f(x^*) + (x - x^*)'\nabla f(x^*)$$

Theorem

Assume f is continuously differentiable on a Set S.

f is convex on S if and only if

$$f(y) \geq f(x) + (y-x)'\nabla f(x) \quad \forall x, y \in S$$

Theorem

- Consider problem min f(x) unconstrained.
- If $\nabla f(\bar{x}) = 0$ and f is convex, then $\bar{x}$ is a global minimum.

Proof:

$\forall y$

$$f(y) \geq f(\bar{x}) + (y - \bar{x})'\nabla f(\bar{x}) \text{ by convexity of } f$$

$$= f(\bar{x}) \text{ since } \nabla f(\bar{x}) = 0.$$

Unconstrained Optimality Conditions

- Basic Problem:

(1) $$\min_{x \in S} f(x)$$

Where S is an open set

e.g. R^n

First Order Necessary Conditions

Theorem: Let f be continuously differentiable.

If x* is a local minimizer of (1),

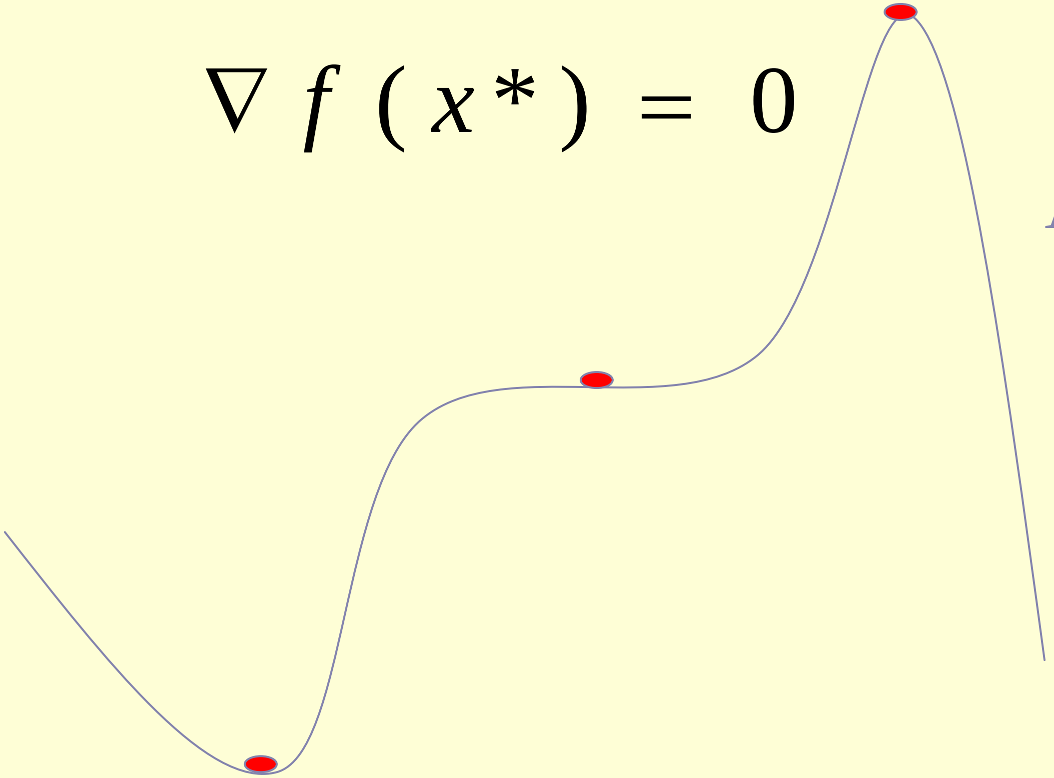
then

$$\nabla f(x^*) = 0$$

Stationary Points

- Note that this condition is not sufficient

$$\nabla f(x^*) = 0$$



Also true for local max and saddle points

Proof

- Assume false, e.g., $\nabla f(x^*) \neq 0$

Let $d = -\nabla f(x^*)$, then

$$f(x^*+\lambda d) = f(x^*) + \lambda d'\nabla f(x^*) + \|\lambda d\| \alpha(x^*, \lambda d)$$

$$\Downarrow$$

$$\frac{f(x^*+\lambda d) - f(x^*)}{\lambda} = d'\nabla f(x^*) + \|d\| \alpha(x^*, \lambda d)$$

$$\Downarrow$$

$f(x^*+\lambda d) - f(x^*) < 0$ for λ sufficiently small

since $d'\nabla f(x^*) < 0$ and $\alpha(x^*, \lambda d) \to 0$.

CONTRADICTION!! x^* is a local min.

Second Order Sufficient Conditions

Theorem: Let f be twice continuously differentiable.

If $\nabla f(x^*) = 0$ and

$$\nabla^2 f(x^*) \text{ is positive definite}$$

then x* is a strict local minimizer of (1).

Proof

- Any point x in neighborhood of x* can be written as $x^*+\lambda d$ for some vector d with norm 1 and $\lambda <= \lambda^*$.

  Since f is twice continuously differentiable, we can choose λ^* such that $\nabla^2 f(\varepsilon)$ is p.d. for all ε such that $\|\varepsilon - x^*\| \le \lambda^*$

$$\forall d, \lambda \le \lambda^*,\ f(x^*+\lambda d) = f(x^*) + \lambda d' \nabla f(x^*) + \frac{1}{2}\lambda^2 d' \nabla^2 f(\varepsilon) d$$

$$\Downarrow$$

since $\nabla f(x^*) = 0$

$$f(x^*+\lambda d) - f(x^*) = \frac{1}{2}\lambda^2 d' \nabla^2 f(x^*) d > 0$$

$$\Downarrow$$

therefore x^* is a strict local min.

Second Order Necessary Conditions

Theorem: Let f be twice continuously differentiable.

If x* is a local minimizer of (1)

then

$$\nabla f(x^*) = 0$$

$$\nabla^2 f(x^*) \text{ is positive semi definite}$$

Proof by contradiction

- Assume false, namely there exists some d such that

$d'\nabla^2 f(x^*)d < 0$ then

$$f(x^*+\lambda d) = f(x^*) + \lambda d'\nabla f(x^*) + \frac{1}{2}\lambda^2 d'\nabla^2 f(x^*)d + \|\lambda d\|^2 \alpha(x^*,\lambda d)$$

$$\Downarrow$$

$$\frac{f(x^*+\lambda d) - f(x^*)}{\lambda^2} = \frac{1}{2} d'\nabla^2 f(x^*)d + \|d\|^2 \alpha(x^*,\lambda d)$$

$$\Downarrow$$

$f(x^*+\lambda d) - f(x^*) < 0$ for λ sufficiently small

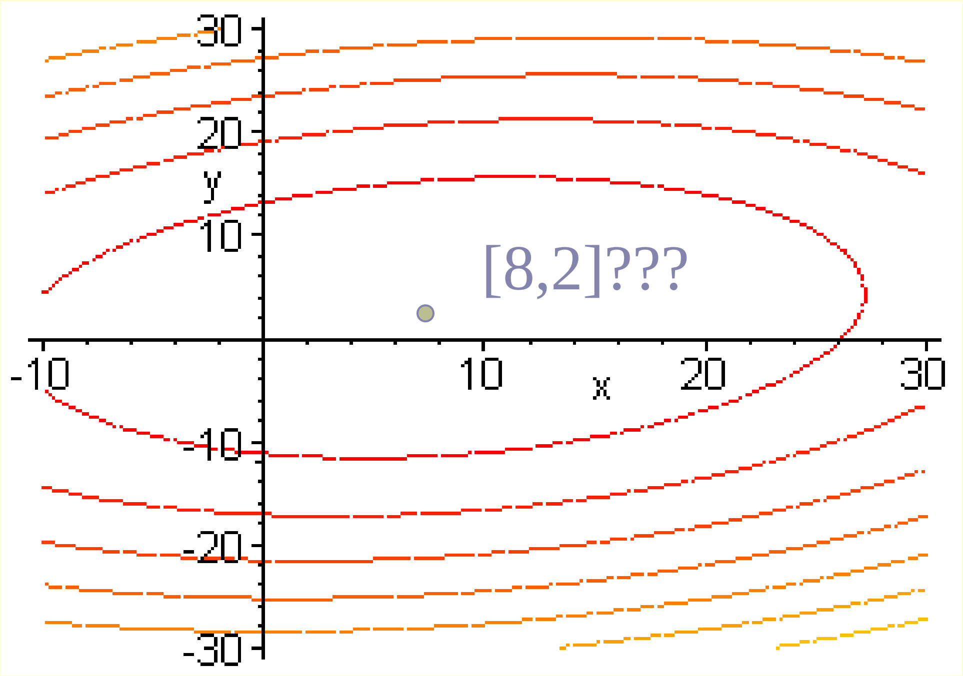
since $d'\nabla f(x^*)d < 0$ for some d and $\alpha(x^*,\lambda d) \to 0$.

Contradiction!!! x* is a local min.

Example

- Say we are minimizing

$$f(x_1, x_2) = x_1^2 - \frac{1}{2}x_1x_2 + 2x_2^2 - 15x_1 - 4x_2$$



Solve FONC

- Solve FONC to find stationary point *

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2 & -\frac{1}{2} \\ -\frac{1}{2} & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 15 \\ 4 \end{bmatrix} = 0$$

$$\Rightarrow \begin{bmatrix} x_1* \\ x_2* \end{bmatrix} = \begin{bmatrix} 2 & -\frac{1}{2} \\ -\frac{1}{2} & 4 \end{bmatrix}^{-1} \begin{bmatrix} 15 \\ 4 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$$

Check SOSC

- The Hessian at x*

$$\nabla^2 f(x_1^*, x_2^*) = \begin{bmatrix} 2 & -\frac{1}{2} \\ -\frac{1}{2} & 4 \end{bmatrix}$$

is p.d. since the eigenvalues,

4.118 and 1.882, are positive.

Therefore SOSC are satisfied.

x* is a strictly local min;

Alternative Argument

- The Hessian at every value x is

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 2 & -\frac{1}{2} \\ -\frac{1}{2} & 4 \end{bmatrix}$$

which is p.d. since the eigenvalues, 4.118 and 1.882, are positive. Therefore the function is strictly convex. Since $\nabla f(x^*)=0$ and f is a strictly convex, x^* is the unique strict global minimum.

You Try

- Use FONC and matlab to find solution of

$$\min f(x_1, x_2, x_3) = 10x_1^2 + 5x_2^2 + 3x_3^2 - 2x_1x_2 + x_2x_3 - 4x_2 - x_1 + x_3$$

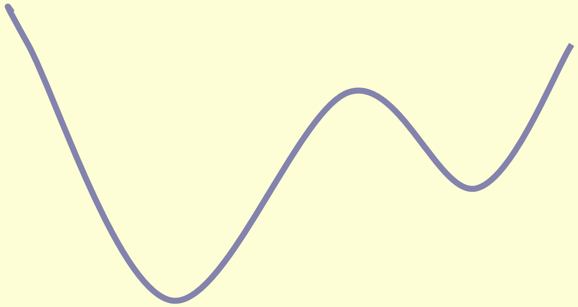
$$\min f(x_1, x_2, x_3) = 10x_1^2 + 5x_2^2 - 3x_3^2 - 2x_1x_2 + x_2x_3 - 4x_2 - x_1 + x_3$$

- Are SOSC satisfied? Are SONC?
- Is f convex?

Optimality Conditions for 1-dimen. functions

- First Order Necessary Condition

If x^* is a local min then $f'(x^*)=0$.



If $f'(x^*)=0$ then ??????????

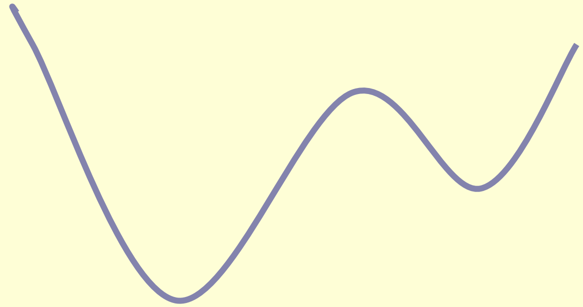
2nd Derivatives - 1D Case

- Sufficient conditions
  - If $f'(x^*)=0$ and $f''(x^*)>0$, then x^* is a strict local min.
  - If $f'(x^*)=0$ and $f''(x^*)<0$, then x^* is a strict local max.
- Necessary conditions
  - If x^* is a local min, then $f'(x^*)=0$ and $f''(x^*)>=0$.
  - If x^* is a local max, then $f'(x^*)=0$ and $f''(x^*)<=0$.

Optimality Conditions for function of R^n

- First Order Necessary Condition

If x* is a local min then $\nabla f(x^*) = 0$



If $\nabla f(x^*) = 0$ then ??????????

Second Order Conditions

- Sufficient conditions
  - If $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is p.d.
    then x* is a strict local min.
  - If $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is n.d.
    then x* is a strict local max.

Second Order Conditions

- Necessary conditions
  - If x* is a local min,

    then $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is p.s.d.
  - If x* is a local max,

    then $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is n.s.d.

Optimality Conditions for Convex Convexity

Let f be continuously differentiable convex function,

- x* is a global minimum of f **if and only if**

$$\nabla f(x^*) = 0$$

Let f be continuously differentiable strictly convex function,

- If $\nabla f(x^*) = 0$ then x* is the unique global minimum of f.

Works similarly for max and concave.

Line Search

- Assume f function maps the vector f to a scalar: $f : R^n \to R$
- Current point is $\overline{x} \in R^n$
- Have interval $\lambda \in [a,b]$
- Want to find:

$$\min_{\lambda \in [a,b]} f(\overline{x} + \lambda d) = g(\lambda)$$

Example

- Say we are minimizing

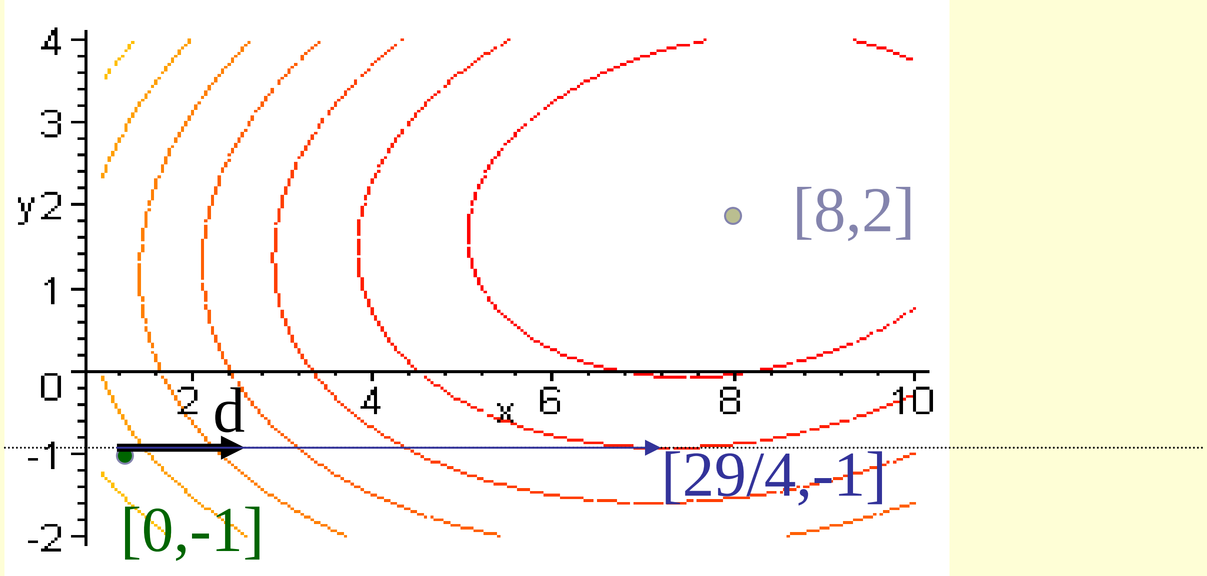
$$f(x_1,x_2)=x_1^2-\frac{1}{2}x_1x_2+2x_2^2-15x_1-4x_2$$

$$=\frac{1}{2}[x_1\ x_2]\begin{bmatrix}2 & -\frac{1}{2}\\ -\frac{1}{2} & 4\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}-[15\quad 4]\begin{bmatrix}x_1\\ x_2\end{bmatrix}$$

- Solution is [8 2]'
- Say we are at [0, -1]' and we want to do a linesearch in direction d=[1 0]'

Line Search

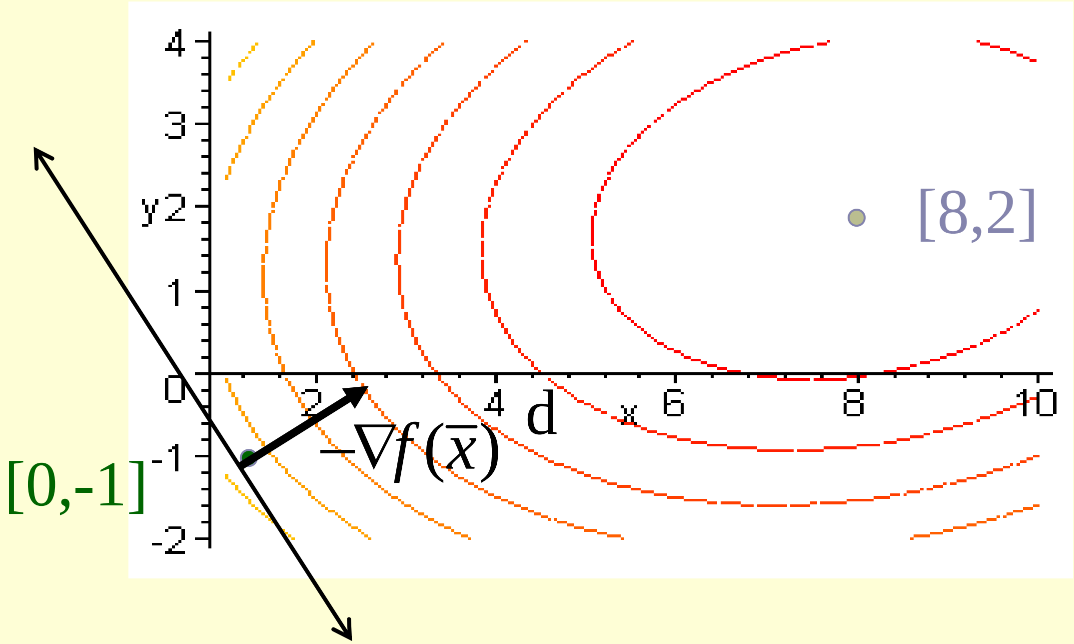
- We are at [0, -1]’ and we want to do a linesearch in direction d=[1 0]’



Descent Directions

- If the directional derivative is negative then linesearch will lead to decrease in the function

$$\nabla f(\bar{x})'d < 0$$



Example continued

- The exact stepsize can be found

$$\bar{x} + \lambda d = \begin{bmatrix} 0 \\ -1 \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \lambda \\ -1 \end{bmatrix}$$

$$g(\lambda) = f(\bar{x} + \lambda d) = \lambda^2 + \frac{\lambda}{2} + 2 - 15\lambda + 4$$

$$g'(\lambda) = f'(\bar{x} + \lambda d) = \nabla f(\bar{x})' d = \left(\begin{bmatrix} 2 & -\frac{1}{2} \\ -\frac{1}{2} & 4 \end{bmatrix} \begin{bmatrix} \lambda \\ -1 \end{bmatrix} - \begin{bmatrix} 15 \\ 4 \end{bmatrix} \right)' \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$= 2\lambda + \frac{1}{2} - 15 = 0$$

$$\Rightarrow \bar{\lambda} = \frac{29}{4}$$

Example continued

- So new point is

$$\overline{x} + \overline{\lambda} d = \begin{bmatrix} 0 \\ -1 \end{bmatrix} + \frac{29}{4}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{29}{4} \\ -1 \end{bmatrix}$$

f([0,-1]=6

f([29/4, -1])=-46.5

f([8,2])=-64

In this case, $f(\overline{x} + \lambda d)$ is a convex function (verify) so this is a Global min.

Example 2

- Consider

$$\min \quad x_1^3 - x_1^2 x_2 + 2x_2^2$$

- Find all points satisfy FONC
- What can you say about that point based on SONC?

FONC

The first order necessary conditions are

$$3x_1^2 - 2x_1x_2 = 0$$

$$-x_1^2 + 4x_2 = 0$$

So both (0,0) and (6,9) satisfy FONC

Second Order Conditions

The Hessian is $\nabla^2 f(x) = \begin{bmatrix} 6x_1 - 2x_2 & -2x_1 \\ -2x_1 & 4 \end{bmatrix}$

The Hessian at (6,9)

$$\nabla^2 f(6,9) = \begin{bmatrix} 18 & -12 \\ -12 & 4 \end{bmatrix}$$

Is indefinite so (6,9) is not a local min (or max).

Second Order Conditions

The Hessian at (0,0)

$$\nabla^2 f(0,0) = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}$$

is psd. So this point satisfies the SONC

But not the SOSC.

It might be a local min, (but in fact it is not try $(-\varepsilon,0)$ for any $\varepsilon>0$

Do Lab 2