Convexity and Optimization

Richard Lusby

DTU Management Engineering

DTU Management Engineering
Department of Management Engineering

Class Exercises From Last Time

Today's Material

- Extrema
- Convex Function
- Convex Sets
- Other Convexity Concepts
- Unconstrained Optimization

Extrema

Problem

$$\max\ f(\boldsymbol{x}) \text{ s.t. } \boldsymbol{x} \in S$$

$$\max \{f(\boldsymbol{x}) : \boldsymbol{x} \in S\}$$

- Global maximum $\boldsymbol{x}^*$

$$f(\boldsymbol{x}^*) \geq f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in S$$

- Local maximum $\boldsymbol{x}^o$:

$$f(\boldsymbol{x}^o) \geq f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \text{ in a neighborhood around } \boldsymbol{x}^o$$

- strict maximum/minimum defined similarly

Weierstrass theorem:

Theorem

A continuous function achieves its max/min on a closed and bounded set

Supremum and Infimum

Supremum

The supremum of a set S having a partial order is the least upper bound of S (if it exists) and is denoted sup S.

Infimum

The infimum of a set S having a partial order is the greatest lower bound of S (if it exists) and is denoted inf S.

- If the extrema are not achieved:
  - max $\to$ sup
  - min $\to$ inf
- Examples
  - $\sup\{2, 3, 4, 5\}$?
  - $\sup\{x \in \mathbb{Q} : x^2 < 2\}$?
  - $\inf\{1/x : x > 0\}$?

Finding Optimal Solutions

Every method for finding and characterizing optimal solutions is based on optimality conditions - either necessary or sufficient

Necessary Condition

A condition $C_1(\boldsymbol{x})$ is necessary if $C_1(\boldsymbol{x}^*)$ is satisfied by every optimal solution $\boldsymbol{x}^*$ (and possibly some other solutions as well).

Sufficient Condition

A condition $C_2(\boldsymbol{x})$ is sufficient if $C_2(\boldsymbol{x}^*)$ ensures that $\boldsymbol{x}^*$ is optimal (but some optimal solutions may not satisfy $C_2(\boldsymbol{x}^*)$.

Mathematically

$$\{\boldsymbol{x}|C_2(\boldsymbol{x})\} \subseteq \{\boldsymbol{x}|\boldsymbol{x} \text{ optimal solution }\} \subseteq \{\boldsymbol{x}|C_1(\boldsymbol{x})\}$$

Finding Optimal Solutions

An example of a necessary condition in the case S is “well-behaved” no improving feasible direction

Feasible Direction

Consider $\boldsymbol{x}^o \in S, s \in \mathbb{R}^n$ is called a feasible direction if there exists $\bar{\epsilon}(s) > 0$ such that

$$\boldsymbol{x}^o + \epsilon s \in S \quad \forall \epsilon : 0 < \epsilon \leq \bar{\epsilon}(s)$$

We denote the cone of feasible directions from $\boldsymbol{x}^o$ in S as $S(x^o)$

Improving Direction

$s \in \mathbb{R}^n$ is called an improving direction if there exists $\bar{\epsilon}(s) > 0$ such that

$$f(\boldsymbol{x}^o + \epsilon s) < f(\boldsymbol{x}^o) \quad \forall \epsilon : 0 < \epsilon \leq \bar{\epsilon}(s)$$

The cone of improving directions from $\boldsymbol{x}^o$ in S is denoted $F(\boldsymbol{x}^o)$

Finding Optimal Solutions

Local Optima

If $\boldsymbol{x}^o$ is a local minimum, then there exist no $s \in S(\boldsymbol{x}^o)$ for which f(.) decreases along s, i.e. for which

$$f(\boldsymbol{x}^o + \epsilon_2 s) < f(\boldsymbol{x}^o + \epsilon_1 s) \text{ for } 0 \leq \epsilon_1 < \epsilon_2 \leq \bar{\epsilon}(s)$$

Stated otherwise: A necessary condition for local optimality is

$$F(\boldsymbol{x}^o) \cap S(\boldsymbol{x}^o) = \emptyset$$

Improving Feasible Directions

If for a given direction s it holds that

$$\nabla f(x^o)s < 0$$

Then s is an improving direction

Well known necessary condition for local optimality of $\boldsymbol{x}^o$ for a differentiable function:

$$\nabla f(\boldsymbol{x}^o) = 0$$

In other words, $\boldsymbol{x}^o$ is stationary with respect to $f(.)$

What if stationarity is not enough?

- Suppose f is twice continuously differentiable
- Analyse the Hessian matrix for f at x^o

$$\nabla^2 f(\boldsymbol{x}^o) = \left\{ \frac{\partial^2 (f(\boldsymbol{x}^o))}{\partial x_i \partial x_j} \right\}, \quad i,j = 1, \dots, n$$

Sufficient Condition

If $\nabla f(\boldsymbol{x}^o) = 0$ and $\nabla^2 f(\boldsymbol{x}^o)$ is positive definite:

$$\boldsymbol{x}^T \nabla^2 f(\boldsymbol{x}^o) \boldsymbol{x} > 0 \quad \forall \quad \boldsymbol{x} \in \mathbb{R}^n \setminus \{0\}$$

then $\boldsymbol{x}^o$ is a local minimum

What if Stationarity is not Enough?

- A necessary condition for local optimality is "Stationarity + positive semidefiniteness of $\nabla^2 f(x^o)$"
- Note that positive definiteness is not a necessary condition
  - E.g. look at $f(x) = x^4$ for $x^o = 0$
- Similar statements hold for maximization problems
  - Key concept here is negative definiteness

Definiteness of a Matrix

- A number of criteria regarding the definiteness of a matrix exist
- A symmetric $n \times n$ matrix A is positive definite if and only if

$$\boldsymbol{x}^T A \boldsymbol{x} > 0 \quad \forall \boldsymbol{x} \in \mathbb{R}^n \backslash \{0\}$$

- Positive semidefinite is defined likewise with $'' \geq ''$ instead of $'' >''$
- Negative (semi) definite is defined by reversing the inequality signs to $'' <''$ and $'' \leq ''$, respectively.

Necessary conditions for positive definiteness:

- A is regular with $det(A) > 0$
- A^{-1} is positive definite

Definiteness of a Matrix

Necessary+Sufficient conditions for positive definiteness:

- Sylvestor's Criterion: All principal submatrices have positive determinants

$$(a_{11}) \quad \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

- All eigenvalues of A are positive

Necessary and Sufficient Conditions

Theorem

Suppose that $f(.)$ is differentiable at a local minimum $\boldsymbol{x}^o$. Then $\nabla f(\boldsymbol{x}^o)s \geq 0$ for $s \in S(\boldsymbol{x}^o)$. If $f(.)$ is twice differentiable at $\boldsymbol{x}^o$ and $\nabla f(\boldsymbol{x}^o) = 0$, then $s^T \nabla^2 f(\boldsymbol{x}^o)s \geq 0 \; \forall s \in S(\boldsymbol{x}^o)$

Theorem

Suppose that S is convex and non-empty, $f(.)$ differentiable, and that $\boldsymbol{x}^o \in S$. Suppose furthermore that $f(.)$ is convex.

- $\boldsymbol{x}^o$ is a local minimum if and only if $\boldsymbol{x}^o$ is a global minimum
- $\boldsymbol{x}^o$ is a local (and hence global) minimum if and only if

$$\nabla f(\boldsymbol{x}^o)(\boldsymbol{x} - \boldsymbol{x}^o) \geq 0 \; \forall \boldsymbol{x} \in S$$

Convex Combination

Convex combination

The convex combination of two points is the line segment between them

$$\alpha_1 \boldsymbol{x_1} + \alpha_2 \boldsymbol{x_2} \text{ for } \alpha_1, \alpha_2 \geq 0 \text{ and } \alpha_1 + \alpha_2 = 1$$

Convex function

Convex Functions

A convex function lies below its chord

$$f(\alpha_1 \boldsymbol{x_1} + \alpha_2 \boldsymbol{x_2}) \leq \alpha_1 f(\boldsymbol{x_1}) + \alpha_2 f(\boldsymbol{x_2})$$

- A strictly convex function has no more than one minimum
- Examples: $y = x^2$, $y = x^4$, $y = x$
- The sum of convex functions is also convex
- A differentiable convex function lies above its tangent
- A differentiable function is convex if its Hessian is positive semi-definite
    - Strictly convex not analogous!
- A function f is concave iff $-f$ is convex

Convex function

Convex Functions

A convex function lies below its chord

$$f(\alpha_1 \boldsymbol{x_1} + \alpha_2 \boldsymbol{x_2}) \leq \alpha_1 f(\boldsymbol{x_1}) + \alpha_2 f(\boldsymbol{x_2})$$

- A strictly convex function has no more than one minimum
- Examples: $y = x^2$, $y = x^4$, $y = x$
- The sum of convex functions is also convex
- A differentiable convex function lies above its tangent
- A differentiable function is convex if its Hessian is positive semi-definite
  - Strictly convex not analogous!
- A function f is concave iff $-f$ is convex

EOQ objective function: $T(Q) = dK/Q + cd + hQ/2$*?*

Economic Order Quantity Model

The problem

The Economic Order Quantity Model is an inventory model that helps manafacturers, retailers, and wholesalers determine how they should optimally replenish their stock levels.

Costs

- K = Setup cost for odering one batch
- c = unit cost for producing/purchasing
- h = holding cost per unit per unit of time in inventory

Assumptions

- d = A known constant demand rate
- Q = The order quantity (arrives all at once)
- Planned shortages are not allowed

Convex sets

Definition

A convex set contains all convex combinations of its elements

$$\alpha_1 \boldsymbol{x_1} + \alpha_2 \boldsymbol{x_2} \in S \; \forall \; \boldsymbol{x_1}, \boldsymbol{x_2} \in S$$

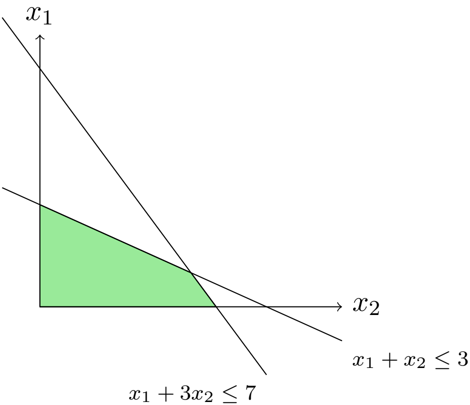
- Some examples of E.g. $(1, 2]$, $x^2 + y^2 < 4$, $\emptyset$
- Level curve (2 dimensions):

$$\{(x, y) : f(x, y) = \beta\}$$

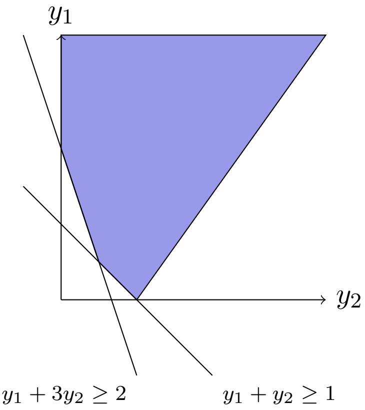
- Level set:

$$\{\boldsymbol{x} : f(\boldsymbol{x}) \leq \beta\}$$

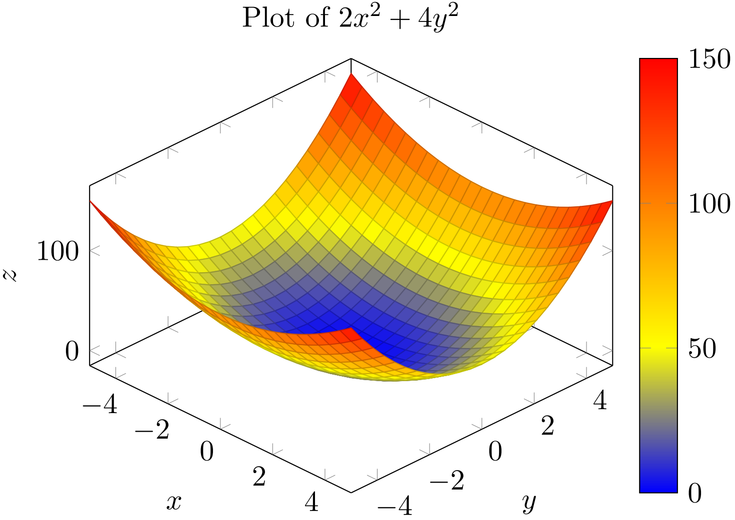
Lower Level Set Example



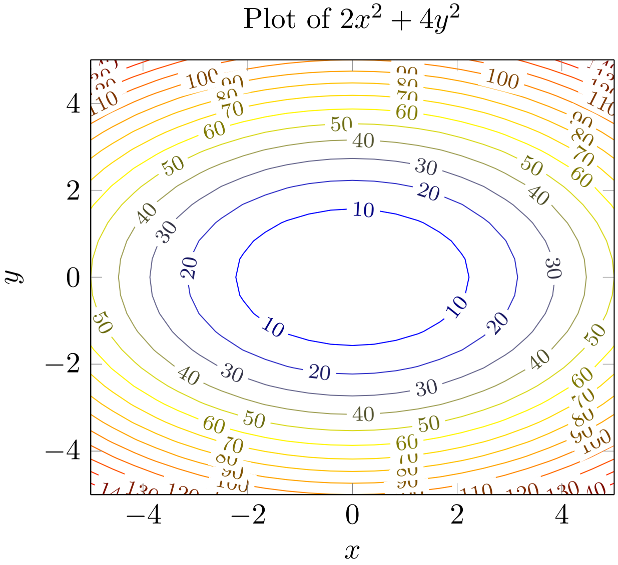
Lower Level Set Example



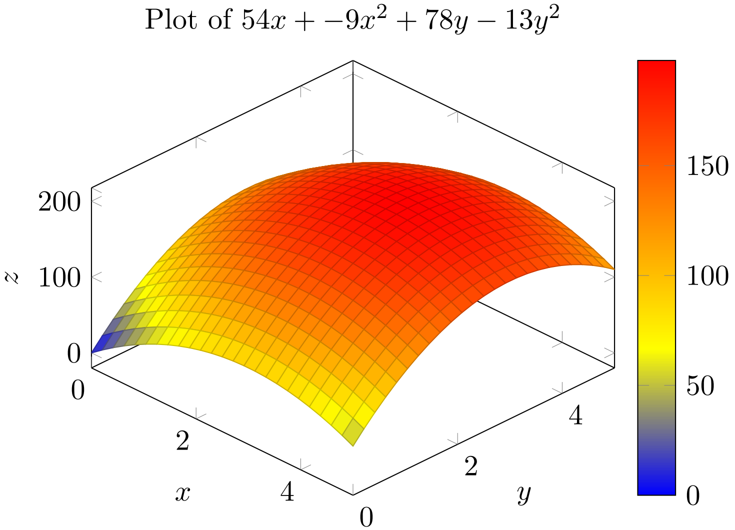
Upper Level Set Example



Upper Level Set Example

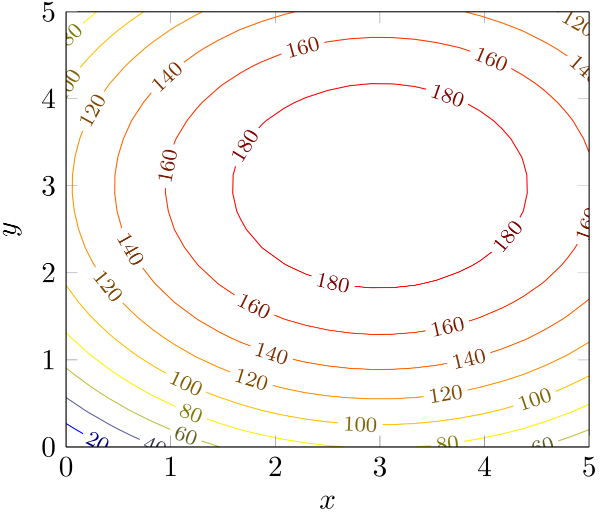


Example



Example

Plot of $54x + -9x^2 + 78y - 13y^2$



Convexity, Concavity, and Optima

Theorem

Suppose that S is convex and that $f(\boldsymbol{x})$ is convex on S for the problem $\min_{\boldsymbol{x} \in S} f(\boldsymbol{x})$, then

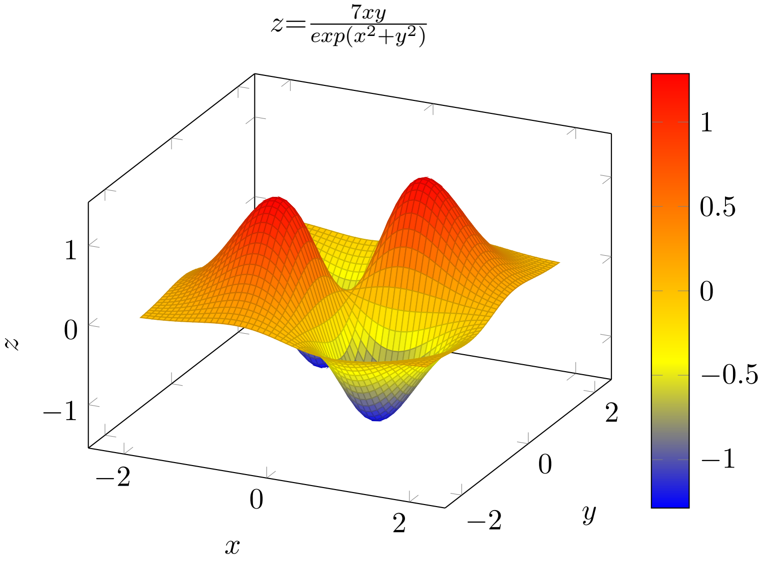
- If $\boldsymbol{x}^*$ is locally minimal, then $\boldsymbol{x}^*$ is globally minimal
- The set X^* of global optimal solutions is convex
- If f is strictly convex, then $\boldsymbol{x}^*$ is unique

Examples

Problem 1

$$\begin{aligned} \text{Minimize} \quad & -x_2 ln x_1 + \tfrac{x_1}{9} + x_2^2 \\ \text{Subject to:} \quad & 1.0 \leq x_1 \leq 5.0 \\ & 0.6 \leq x_2 \leq 3.6 \end{aligned}$$

What does the function look like?



Problem 2

$$\begin{array}{lrl} \text{Minimize} & \sum_{i=1}^{3} -i ln x_i & \\ \text{Subject to:} & \sum_{i=1}^{3} x_i & = 6 \\ & x_i & \leq 3.5 \; i = 1,2,3 \\ & x_i & \geq 1.5 \; i = 1,2,3 \end{array}$$

Class exercises

- Show that $f(\boldsymbol{x}) = ||\boldsymbol{x}|| = \sqrt{\sum_i x_i^2}$ is convex
- Prove that any level set of a convex function is a convex set

Other Types of Convexity

- The idea of pseudoconvexity of a function is to extend the class of functions for which stationarity is a sufficient condition for global optimality. If f is defined on an open set X and is differentiable we define the concept of pseudoconvexity.
- A differentiable function f is pseudoconvex if

$$\nabla f(\boldsymbol{x}) \cdot (\boldsymbol{x'} - \boldsymbol{x}) \geq 0 \Rightarrow f(\boldsymbol{x'}) \geq f(\boldsymbol{x}) \quad \forall x, x' \in X$$

- or alternatively ..

$$f(\boldsymbol{x}') < f(\boldsymbol{x}) \Rightarrow \nabla f(\boldsymbol{x})(\boldsymbol{x}' - \boldsymbol{x}) < 0 \quad \forall x, x' \in X$$

- A function f is pseudoconcave iff $-f$ is pseudoconvex
- Note that if f is convex and differentiable, and X is open, then f is also pseudoconvex

Other Types of Convexity

- A function is quasiconvex if all lower level sets are convex
- That is, the following sets are convex

$$S' = \{\boldsymbol{x} : f(\boldsymbol{x}) \leq \beta\}$$

- A function is quasiconcave if all upper level sets are convex
- That is, the following sets are convex

$$S' = \{\boldsymbol{x} : f(\boldsymbol{x}) \geq \beta\}$$

- Note that if f is convex and differentiable, and X is open, then f is also quasiconvex
- Convexity properties

Convex $\Rightarrow$ pseudoconvex $\Rightarrow$ quasiconvex

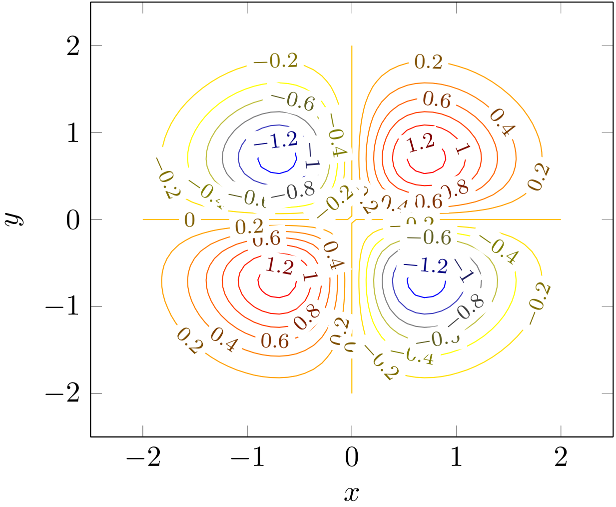
Exercises

Show the following

- $f(x) = x + x^3$ is pseudoconvex but not convex
- $f(x) = x^3$ is quasiconvex but not pseudoconvex

Graphically

$$z = \frac{7xy}{exp(x^2+y^2)}$$



Exercises

Convexity Questions

- Can a function be both convex and concave?
- Is a convex function of a convex function convex?
- Is a convex combination of convex functions convex?
- Is the intersection of convex sets convex?

Unconstrained problem

$$\min f(\boldsymbol{x}) \text{ s.t. } \boldsymbol{x} \in R^n$$

- Necessary optimality condition for $\boldsymbol{x^o}$ to be a local minimum

$$\nabla f(\boldsymbol{x}^o) = 0 \text{ and } H(\boldsymbol{x}^o) \text{ is positive semidefinite}$$

- Sufficient optimality condition for $\boldsymbol{x^o}$ to be a local minimum

$$\nabla f(\boldsymbol{x}^o) = 0 \text{ and } H(\boldsymbol{x}^o) \text{ is positive definite}$$

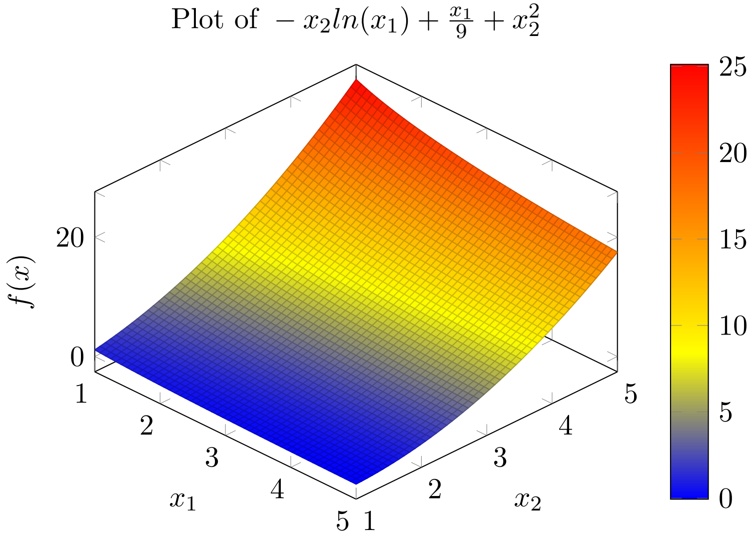
- Necessary and sufficient
  - Suppose f is pseudoconvex
  - $\boldsymbol{x}^*$ is a global minimum iff $\nabla f(\boldsymbol{x}^*) = 0$

Unconstrained example

$$\min f(x) = (x^2 - 1)^3$$

- $f'(x) = 6x(x^2 - 1)^2 = 0$ for $x = 0, \pm 1$
- $H(x) = 24x^2(x^2 - 1) + 6(x^2 - 1)^2$
- $H(0) = 6$ and $H(\pm 1) = 0$
- Therefore $x = 0$ is a local minimum (actually the global minimum)
- $x = \pm 1$ are saddle points

What does the function look like?



Class Exercise

Problem

Suppose A is an $m * n$ matrix, $\boldsymbol{b}$ is a given m vector, find

$$\min ||A\boldsymbol{x} - \boldsymbol{b}||^2$$

Richard M. Lusby
DTU Management Engineering, Technical University of Denmark

Building 424, Room 208
2800 Kgs. Lyngby, Denmark
http://www.man.dtu.dk

rmlu@dtu.dk
phone +45 4525 3084