

3 Multidimensional search

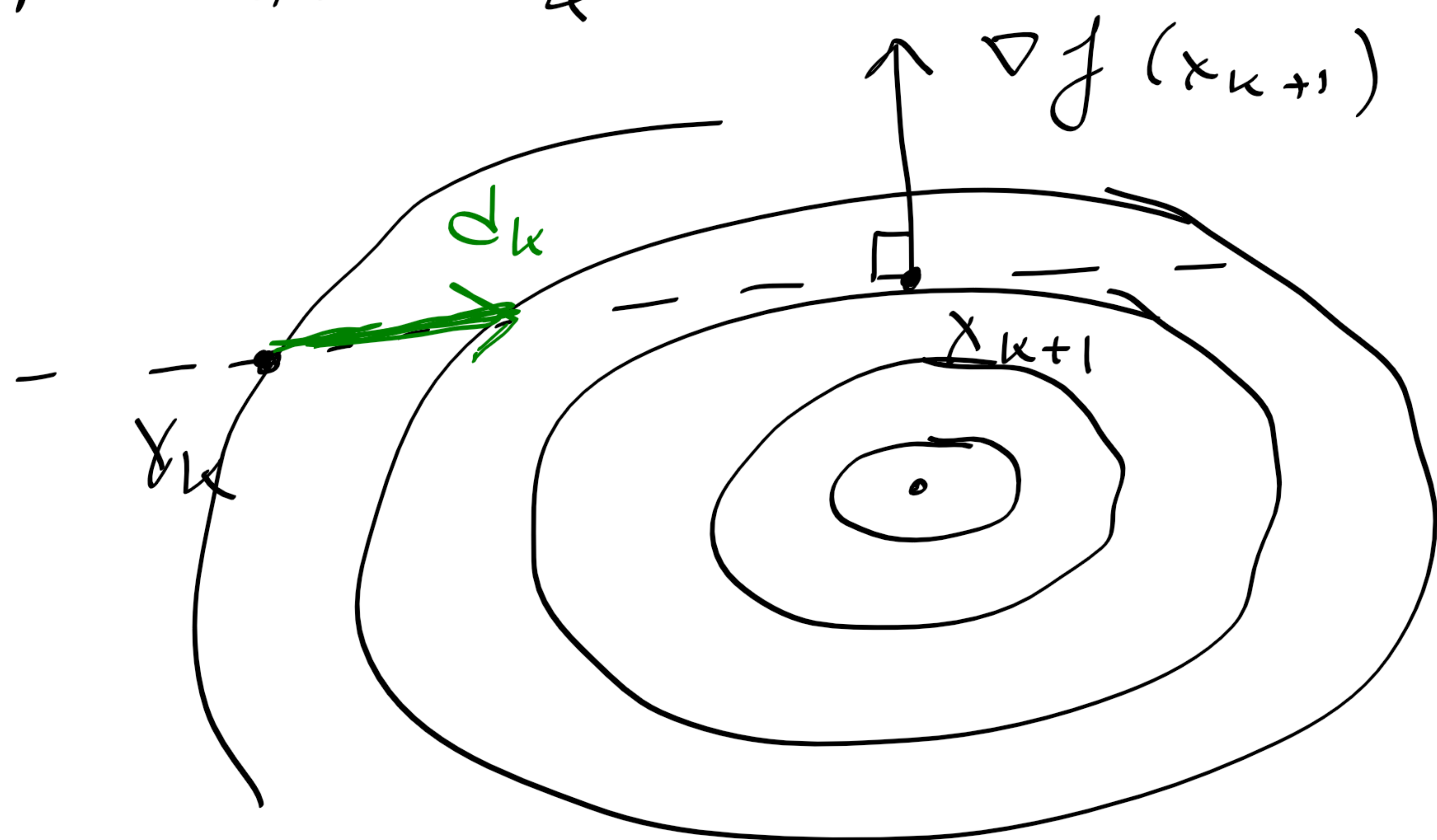
$$(P) \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x)$$

Pradtype algorithm

- given x_k , choose search direction d_k
- line search: minimize $F(\lambda) = f(x_k + \lambda d_k)$
exact or inexact to define λ_k
- set $x_{k+1} = x_k + \lambda_k d_k$

Lemma 1: exact line search for $f \in C^1$ implies that $\nabla f(x_{k+1})^T d_k = 0$.

$$\text{Proof: } F'(\lambda) = \nabla f(x_k + \lambda d_k)^T d_k \quad \Rightarrow$$
$$0 = F'(\lambda_k) = \nabla f(x_{k+1})^T d_k \quad \neq$$



Cartagena99

CLASES PARTICULARES, TUTORIAS TÉCNICAS ONLINE

LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS

CALL OR WHATSAPP: 689 45 44 70

Exact line search for a quadratic function

$$q(x) = \frac{1}{2} x^T H x + c^T x \quad \text{with } H \text{ pos. def.}$$

$$d_k^T \nabla q(x_{k+1}) = 0 \quad (\text{Lemma 1}) \quad \Leftrightarrow$$

$$d_k^T (H x_{k+1} + c) = 0 \quad \Leftrightarrow$$

$$d_k^T (H (x_k + \lambda_k d_k) + c) = 0 \quad (=)$$

$$d_k^T H x_k + \lambda_k d_k^T H d_k + d_k^T c = 0 \quad (=)$$

$$\lambda_k = - \frac{d_k^T (H x_k + c)}{d_k^T H d_k} = - \frac{d_k^T \nabla q(x_k)}{d_k^T H d_k}$$

Cartagena99

CLASES PARTICULARES, TUTORIAS TÉCNICAS ONLINE

LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS

CALL OR WHATSAPP: 689 45 44 70

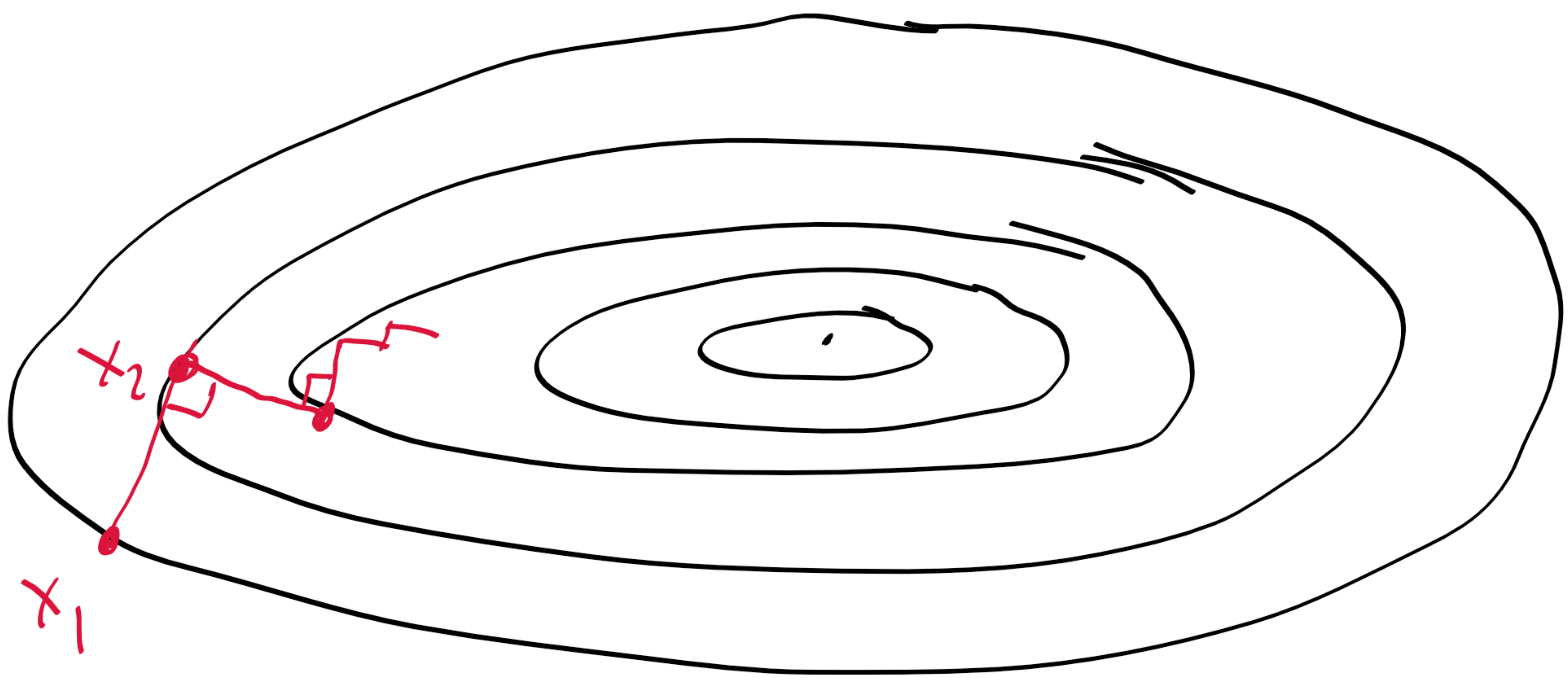
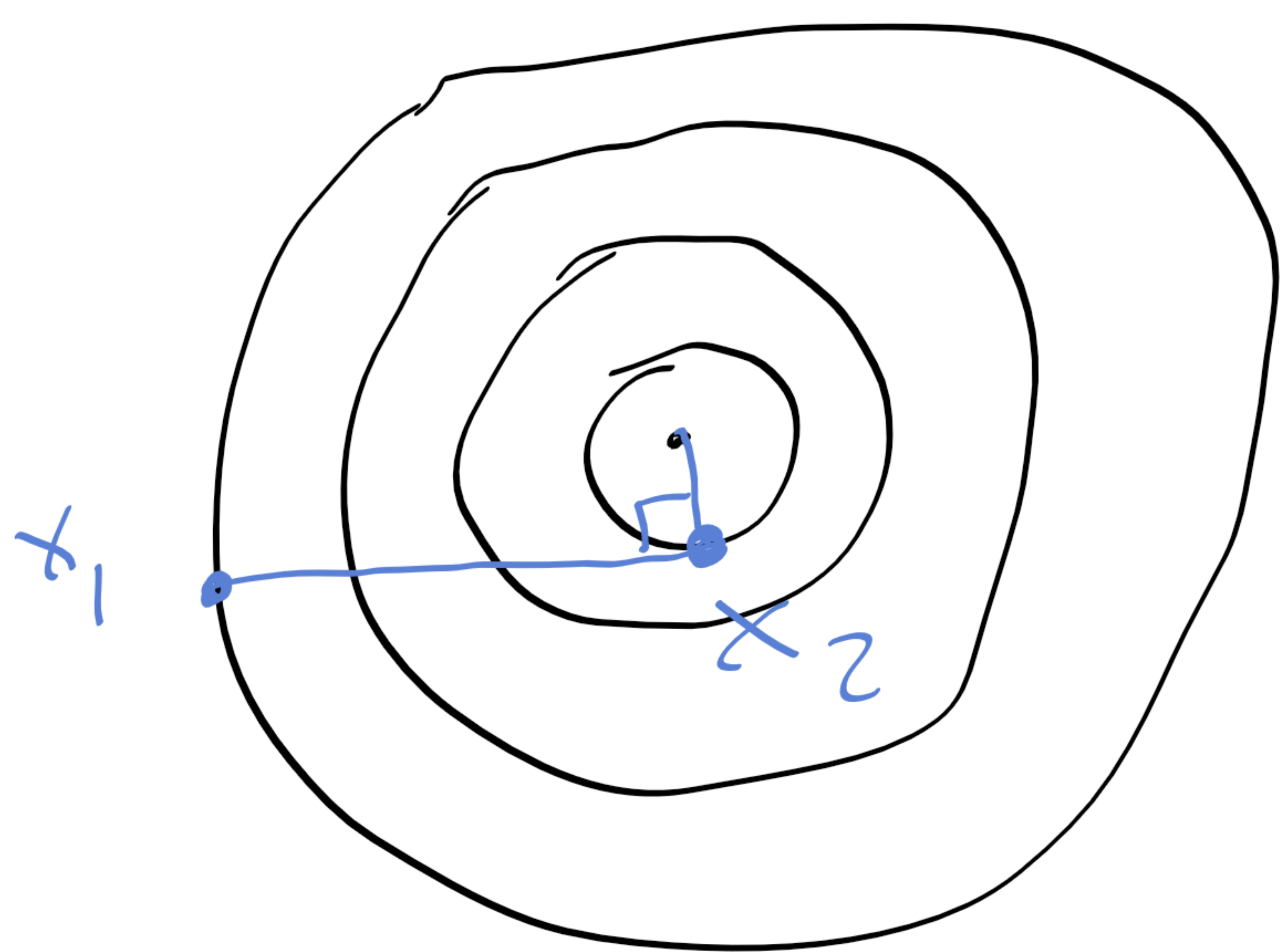
3.3 Steepest descent

search direction $d_k = -\nabla f(x_k)$

with exact line search

$$0 = \nabla f(x_{k+1})^T d_k = -d_{k+1}^T d_k$$

Thus, a zig-zagging behaviour:



Property on a quadratic function with Hessian H with eigenvalues $0 < \lambda_{\min} \leq \lambda_2 \leq \dots \leq \lambda_{\max}$ and exact line search:

Then 1: $|g(x_{k+1}) - g_{\min}| \leq C |g(x_k) - g_{\min}|$ with

$$C = \left(\frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} \right)^2 = \left(\frac{\alpha - 1}{\alpha + 1} \right)^2 < 1$$

where the condition number $\alpha = \frac{\lambda_{\max}}{\lambda_{\min}}$

Fast convergence if $\alpha \approx 1$

Slow convergence if $\alpha \gg 1$

Cartagena99

CLASES PARTICULARES, TUTORIAS TÉCNICAS ONLINE

LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS

CALL OR WHATSAPP: 689 45 44 70

Appendix A. Some matrix theory

THEOREM 1. Let \mathbf{A} be a quadratic $n \times n$ matrix. The following conditions on \mathbf{A} are equivalent.

- The columns of \mathbf{A} form a basis for \mathbb{R}^n .
- The rows of \mathbf{A} form a basis for \mathbb{R}^n .
- The rank of \mathbf{A} is equal to n .
- \mathbf{A} is invertible.
- The homogeneous system of equations $\mathbf{Ax} = \mathbf{0}$ has the zero solution only.
- The system of equations $\mathbf{Ax} = \mathbf{b}$ has a unique solution for any \mathbf{b} .
- $\det \mathbf{A} \neq 0$.

The spectral theorem for symmetric matrices:

THEOREM 2. Let \mathbf{H} be a symmetric matrix. Then

- there exists an orthogonal basis of eigenvectors of \mathbf{H} ,
- there is an orthogonal matrix \mathbf{Q} and a diagonal matrix Λ such that

$$\mathbf{Q}^{-1}\mathbf{H}\mathbf{Q} = \mathbf{Q}^T\mathbf{H}\mathbf{Q} = \Lambda.$$

In \mathbf{Q} the columns are eigenvectors of \mathbf{H} , in Λ the diagonal elements are the eigenvalues.

A.1 Positive definite matrices

DEFINITION 1. Let \mathbf{H} be a symmetric $n \times n$ matrix. \mathbf{H} is called²

positive definite if $\mathbf{x}^T\mathbf{H}\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$,

positive semidefinite if $\mathbf{x}^T\mathbf{H}\mathbf{x} \geq 0$ for all \mathbf{x} ,

indefinite if $\mathbf{x}^T\mathbf{H}\mathbf{x} > 0$ for some $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{x}^T\mathbf{H}\mathbf{x} < 0$ for

Cartagena99

CLASES PARTICULARES, TUTORIAS TÉCNICAS ONLINE

LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS

CALL OR WHATSAPP: 689 45 44 70

Proof: Take

$\mathbf{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ ← position i . Then $0 < \mathbf{e}_i^T \mathbf{H} \mathbf{e}_i = h_{ii} \neq$

Thm 3: H pos. def. \Leftrightarrow eigenvalues $\lambda_i > 0, i=1, \dots, n$

H semidef $\Leftrightarrow \dots \lambda_i \geq 0 \dots$

Proof: Given $x^T H x$, change variables according to the spectral theorem: $x = Q \hat{x}$ with Q orthogonal. Then

$$x^T H x = \hat{x}^T Q^T H Q \hat{x} = \hat{x}^T \Lambda \hat{x} = \sum_{i=1}^n \lambda_i \hat{x}_i^2 \geq 0$$

Thus pos. semidef. proved.

H pos. def. $\Leftrightarrow x^T H x > 0 \quad \forall x \neq 0 \Leftrightarrow$

$\lambda_i > 0 \quad \forall \hat{x} \neq 0 \Leftrightarrow$

$\lambda_i > 0 \quad \forall Q^T x \neq 0 \Leftrightarrow$

$\lambda_i > 0 \quad \forall x \neq 0 \quad \#$

Cor. H pos. def. $\Rightarrow \det H = \lambda_1 \lambda_2 \dots \lambda_n > 0$

Let A_k denote the upper left $k \times k$ submatrix of A :

$$A = \begin{pmatrix} & & & & \\ & & & & \\ & & A_k & & \\ & & & & \\ & & & & \end{pmatrix}$$

Lemma 1: H pos. def. $\Rightarrow \det H_k > 0, k=1, \dots, n$

Proof: Given $x^T H x > 0 \quad \forall x \neq 0$. Take $|\hat{x}| \leftarrow$ size $k \times 1$

Cartagena99

CLASES PARTICULARES, TUTORIAS TÉCNICAS ONLINE

LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS

CALL OR WHATSAPP: 689 45 44 70

Appendix A.2 LU factorization

Gaussian elimination:

$$Ax = b \iff \begin{pmatrix} * & \dots & * \\ 0 & * & \\ \vdots & & \\ 0 & \dots & 0 & * \end{pmatrix} x = \tilde{b}$$

U upper triangular

If the system needs to be solved many times for different b , then we keep track of row operations in a matrix L :

$$Ax = b \iff LUx = b \iff \begin{cases} L\tilde{b} = b \\ Ux = \tilde{b} \end{cases}$$

The logo for 'Cartagena99' features the text 'Cartagena99' in a stylized, teal-colored font. The '99' is significantly larger and more prominent than the 'Cartagena' part. The text is set against a background of a light blue triangle pointing downwards and an orange triangle pointing upwards, both with a slight gradient and shadow effect.

CLASES PARTICULARES, TUTORIAS TÉCNICAS ONLINE

LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS

CALL OR WHATSAPP: 689 45 44 70

THEOREM 4. Assume that $\det \mathbf{A}_k \neq 0$ for all k .

1. There is a unique factorization $\mathbf{A} = \mathbf{L}\mathbf{U}$, with \mathbf{L} a lower triangular matrix with all diagonal elements equal to 1, and \mathbf{U} an upper triangular matrix;

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ l_{21} & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & 1 \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} d_1 & u_{12} & \cdots & u_{1n} \\ 0 & d_2 & \cdots & u_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}.$$

2. When using Gaussian elimination to solve the linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$ these matrices appear as follows:

- The columns l_1, \dots, l_n of \mathbf{L} contain the multipliers needed in the elimination; specifically, column l_j contains those needed to eliminate x_j from rows $j+1, \dots, n$.
- The result of the eliminations is \mathbf{U} .
- The diagonal elements in \mathbf{U} (the **pivot elements**) are

$$(1) \quad d_k = \frac{\det \mathbf{A}_k}{\det \mathbf{A}_{k-1}}, \quad k = 1, \dots, n$$

(where $\det \mathbf{A}_0$ should be interpreted as 1).

3. In the special case of \mathbf{A} being symmetric, $\mathbf{U} = \mathbf{D}\mathbf{L}^T$, where \mathbf{D} is the diagonal matrix whose diagonal entries are d_1, \dots, d_n . Hence

$$\mathbf{A} = \mathbf{L}\mathbf{D}\mathbf{L}^T$$

when \mathbf{A} is symmetric.

Cartagena99

CLASES PARTICULARES, TUTORIAS TÉCNICAS ONLINE

LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS

CALL OR WHATSAPP: 689 45 44 70

$$A = \begin{pmatrix} 2 & 3 & 4 \\ 4 & 9 & 7 \\ 2 & -3 & 5 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 3 & 4 \\ 0 & 3 & -1 \\ 0 & -6 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 3 & 4 \\ 0 & 3 & -1 \\ 0 & 0 & -1 \end{pmatrix} = U$$

row operations

$$\begin{aligned} r_1 &:= r_1 \\ r_2 &:= r_2 - 2r_1 \\ r_3 &:= r_3 - 1r_1 \end{aligned}$$

$$\begin{aligned} r_1 &:= r_1 \\ r_2 &:= r_2 \\ r_3 &:= r_3 - (-2)r_2 \end{aligned}$$

Alternatively

$$A = \begin{pmatrix} 2 & 3 & 4 \\ 4 & 9 & 7 \\ 2 & -3 & 5 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & -6 & -1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$$\begin{aligned} r_1 &:= r_1 - r_1 \\ r_2 &:= r_2 - 2r_1 \\ r_3 &:= r_3 - 1r_1 \end{aligned}$$

$$\text{set } l_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

$$\begin{aligned} r_1 &:= r_1 - 0r_2 \\ r_2 &:= r_2 - r_2 \\ r_3 &:= r_3 - (-2)r_2 \end{aligned}$$

$$l_2 = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$$

$$u_1^T = (2 \ 3 \ 4)$$

$$u_2^T = (0 \ 3 \ -1)$$

$$l_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad u_3^T = (0 \ 0 \ -1)$$

$$A - l_1 u_1^T - l_2 u_2^T - l_3 u_3^T = 0 \quad \Leftrightarrow$$

$$A - (l_1 \ l_2 \ l_3) \begin{pmatrix} u_1^T \\ u_2^T \\ u_3^T \end{pmatrix} = 0 \quad \Leftrightarrow \quad A - LU = 0$$

This can only be done if all pivot elements but the last one are nonzero:

$$\begin{aligned} d_1 &= 2 \\ d_2 &= 3 \\ d_3 &= -1 \end{aligned}$$

We have $\det A_i = d_i$

Since row operations don't change the

Cartagena99

CLASES PARTICULARES, TUTORIAS TÉCNICAS ONLINE

LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS

CALL OR WHATSAPP: 689 45 44 70

We can compose

$$U = \begin{pmatrix} d_1 & & & \\ & \ddots & & \\ & & \ddots & \\ 0 & & & d_n \end{pmatrix} = \begin{pmatrix} d_1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & d_n \end{pmatrix} \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix} = DM$$

If A is symmetric, then we use uniqueness of LU factorization (see book):

$$LDM = \underbrace{L}_{\text{lower}} U = A = A^T = \underbrace{M^T}_{\text{lower}} \underbrace{D}_{\text{upper}} L^T \Rightarrow L = M^T$$

So $A = LDL^T$ #
(we may still have $d_n = 0$)

The logo for Cartagena99 features the text "Cartagena99" in a stylized, teal-colored font. The "99" is significantly larger and more prominent than the word "Cartagena". The text is set against a light blue background with a white, cloud-like shape behind it. Below the text is a thick, orange-to-yellow gradient bar that tapers at the ends.

CLASES PARTICULARES, TUTORIAS TÉCNICAS ONLINE

LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS

CALL OR WHATSAPP: 689 45 44 70

Appendix A.3

Thm 6. H pos. def. $(\Leftrightarrow) H = LDL^T$ with

L lower triangular with ones on the diagonal and
 $D = \text{diag}(d_1, \dots, d_n)$ with $d_i > 0 \quad \forall i$

Proof: (\Rightarrow) Lemma 1 and Thm 4.

(\Leftarrow) Take an arbitrary $x \neq 0$

$$x^T H x = x^T L D L^T x = \underbrace{(L^T x)^T}_{\hat{x}} D (L^T x) = \hat{x}^T D \hat{x}$$

$$= \sum d_i \hat{x}_i^2 > 0 \quad \text{iff} \quad \hat{x} \quad \text{all } d_i > 0 \quad \#$$

This is related to completing the squares:

$$\hat{x} = L^T x = \begin{pmatrix} 1 & & 0 \\ l_{21} & \ddots & \\ \vdots & \ddots & \\ l_{n1} & \dots & 1 \end{pmatrix}^T x = \begin{pmatrix} 1 & l_{21} & \dots & l_{n1} \\ & 1 & \ddots & \\ & & \ddots & 1 \\ 0 & & & 1 \end{pmatrix} x = \begin{pmatrix} x_1 + l_{21}x_2 + \dots + l_{n1}x_n \\ \vdots \\ x_n \end{pmatrix}$$

$$\hat{x}_1^2 = (x_1 + l_{21}x_2 + \dots + l_{n1}x_n)^2$$

See Ex. 4.

Thm 7 (Sylvester's criterion)

H pos. def. $(\Leftrightarrow) \det H_k > 0, \quad k = 1, \dots, n$

Proof: (\Rightarrow) Lemma 1 (\Leftarrow) Thm 4 gives

$$H = LDL^T \quad \text{with} \quad d_k = \frac{\det H_k}{\det H_{k-1}} > 0$$

$\Rightarrow H$ pos. def. by Thm 6. #

Cartagena99

CLASES PARTICULARES, TUTORIAS TÉCNICAS ONLINE

LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS

CALL OR WHATSAPP: 689 45 44 70

converse is not true

$$\text{Thm 8. } \begin{cases} \det H_k > 0, & k=1, \dots, n-1 \\ \det H_n = \det H = 0 \end{cases}$$

$\Rightarrow H$ pos. semidef.

Proof: Thm 4 $\Rightarrow d_1, \dots, d_{n-1} > 0$ and $d_n = 0$.

Use Thm 6 to conclude the statement. $\#$

Thm 9: (Cholesky factorization)

$$H \text{ pos. def. } \Leftrightarrow H = \hat{L} \hat{L}^T \text{ where}$$

\hat{L} is lower triangular with positive diagonal elements

Proof: (\Rightarrow) $H = L D L^T$ with $L = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ & & 1 \end{pmatrix}$, $D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$
with all $d_i > 0$. Define $\sqrt{D} = \begin{pmatrix} \sqrt{d_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{d_n} \end{pmatrix}$. Then

$$H = L D L^T = L \underbrace{\sqrt{D} \sqrt{D}}_{\hat{L}} L^T = \hat{L} \hat{L}^T$$

(\Leftarrow) Exercise!

Cartagena99

CLASES PARTICULARES, TUTORIAS TÉCNICAS ONLINE

LLAMA O ENVÍA WHATSAPP: 689 45 44 70

ONLINE PRIVATE LESSONS FOR SCIENCE STUDENTS

CALL OR WHATSAPP: 689 45 44 70