Chapter 29

Mathematical background

ap,math

Contents

s,math,intro

29.1 Introduction (s,math,intro)

This appendix reviews some basic mathematical tools and notation used in this book.

s,math,four

29.2 Fourier transforms (s,math,four)

The $\bar{d}$-dimensional Fourier transform $F(\vec{\nu})$ of a function $f(\vec{x})$, where $\vec{x} \in \mathbb{R}^{\bar{d}}$ and $\vec{\nu} \in \mathbb{R}^{\bar{d}}$, and the inverse Fourier transform are given by:

$$F(\vec{\nu}) = \int f(\vec{x})\, \mathrm{e}^{-\imath 2\pi \vec{\nu} \cdot \vec{x}}\, \mathrm{d}\vec{x} \tag{29.2.1}$$

e,math,four,ft

$$f(\vec{x}) = \int F(\vec{\nu})\, \mathrm{e}^{\imath 2\pi \vec{\nu} \cdot \vec{x}}\, \mathrm{d}\vec{\nu}\,. \tag{29.2.2}$$

29.2.1 Properties

29.2.1.1 Linearity

$$\sum_j \alpha_j\, f_j(\vec{x}) \overset{\mathrm{FT}}{\longleftrightarrow} \sum_j \alpha_j\, F_j(\vec{\nu})$$

29.2.1.2 Convolution property

The convolution of two $\bar{d}$-dimensional signals $f(\vec{x})$ and $g(\vec{x})$ is denoted

e,math,four,conv

$$(f * g)(\vec{x}) = \int_{\mathbb{R}^{\bar{d}}} f(\vec{x}')\, g(\vec{x} - \vec{x}')\, \mathrm{d}\vec{x}'\,. \tag{29.2.3}$$

The **convolution property** of the **Fourier transform** is:

e,math,four,conv

$$h(\vec{x}) = (f * g)(\vec{x}) \overset{\mathrm{FT}}{\longleftrightarrow} H(\vec{\nu}) = F(\vec{\nu})\, G(\vec{\nu})\,. \tag{29.2.4}$$

29.2.1.3 Shift property

$$g(\vec{x}) = f(\vec{x} - \vec{x}_0) \overset{\mathrm{FT}}{\longleftrightarrow} G(\vec{\nu}) = \mathrm{e}^{-\imath 2\pi \vec{\nu} \cdot \vec{x}_0}\, F(\vec{\nu})\,.$$

29.2.1.4 Scaling property

$$g(\vec{x}) = f(\alpha \vec{x}) \overset{\mathrm{FT}}{\longleftrightarrow} G(\vec{\nu}) = \frac{1}{|\alpha|^{\bar{d}}}\, F\left(\frac{\vec{\nu}}{\alpha}\right), \qquad \alpha \neq 0.$$

29.2.1.5 Affine scaling property

For an invertible $\bar{d} \times \bar{d}$ matrix $\boldsymbol{B}$: (see Problem 29.1):

$$g(\vec{\mathrm{x}}) = f\left(\boldsymbol{B}^{-1}\left(\vec{\mathrm{x}} - \vec{\mathrm{x}}_0\right)\right) \overset{\mathrm{FT}}{\longleftrightarrow} G(\vec{\nu}) = |\mathsf{det}\{\boldsymbol{B}\}|\, \mathrm{e}^{-\imath 2\pi \vec{\nu} \cdot \vec{\mathrm{x}}_0}\, F(\boldsymbol{B}'\vec{\nu})\,. \tag{29.2.5}$$

e,math,four,affine

29.2.1.6 Rotation property

For an orthonormal $\bar{d} \times \bar{d}$ matrix $\boldsymbol{U}$:

$$g(\vec{\mathrm{x}}) = f(\boldsymbol{U}\vec{\mathrm{x}}) \overset{\mathrm{FT}}{\longleftrightarrow} G(\vec{\nu}) = F(\boldsymbol{U}\vec{\nu})\,. \tag{29.2.6}$$

e,math,four,rotate

29.2.1.7 Parseval's theorem

$$\int f(\vec{\mathrm{x}})\, g^*(\vec{\mathrm{x}})\, \mathrm{d}\vec{\mathrm{x}} = \int F(\vec{\nu})\, G^*(\vec{\nu})\, \mathrm{d}\vec{\nu}\,. \tag{29.2.7}$$

e,math,four,parseval

29.2.1.8 Differentiation property

$$g(\vec{\mathrm{x}}) = \frac{\partial}{\partial x_j} f(\vec{\mathrm{x}}) \overset{\mathrm{FT}}{\longleftrightarrow} G(\vec{\nu}) = \imath 2\pi \nu_j\, F(\vec{\nu})\,.$$

29.2.1.9 Laplacian property

$$g(\vec{\mathrm{x}}) = \Delta^2 f(\vec{\mathrm{x}}) = \sum_{j=1}^{\bar{d}} \frac{\partial^2}{\partial x_j^2} f(\vec{\mathrm{x}}) \overset{\mathrm{FT}}{\longleftrightarrow} G(\vec{\nu}) = \sum_{j=1}^{\bar{d}} -(2\pi\nu_j)^2\, F(\vec{\nu}) = -(2\pi)^2 \left\|\vec{\nu}\right\|^2 F(\vec{\nu})\,. \tag{29.2.8}$$

e,math,four,laplace

29.2.1.10 Circular symmetry

If $f(\vec{\mathrm{x}}) = f_0(\|\vec{\mathrm{x}}\|)$ then $F(\vec{\nu}) = F_0(\|\vec{\nu}\|)$.

29.2.1.11 Separability

If $f(\vec{\mathrm{x}}) = \prod_{i=1}^{\bar{d}} f_i(x_i)$ then $F(\vec{\nu}) = \prod_{i=1}^{\bar{d}} F_i(\nu_i)$, where $f_i(x) \overset{\mathrm{FT}}{\longleftrightarrow} F_i(\nu)$.

29.2.1.12 Hermitian symmetry property

If $f(\vec{\mathrm{x}})$ is **real**, then its Fourier transform is **Hermitian symmetric**: $F(-\vec{\nu}) = F^*(\vec{\nu})$.

29.3 Discrete Fourier transform (s,math,dsp)

s,math,dsp

The **1D N-point discrete Fourier transform** (**DFT**) of a signal $x[n]$ is defined by

$$X[k] = \sum_{n=0}^{N-1} x[n]\, \mathrm{e}^{-\imath \frac{2\pi}{N} nk}\,, \quad k = 0, \ldots, N-1, \tag{29.3.1}$$

e,math,dsp,dft

and the inverse DFT is given by

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k]\, \mathrm{e}^{\imath \frac{2\pi}{N} nk}\,, \quad n = 0, \ldots, N-1. \tag{29.3.2}$$

e,math,dsp,idft

29.3.1 Properties of DFT

- **Shift (periodic)**

$$x[(n - n_0) \bmod N] \overset{\mathrm{DFT}}{\longleftrightarrow} \mathrm{e}^{-\imath \frac{2\pi}{N} n_0}\, X[k]\,.$$

- **Parseval's theorem**

$$\sum_{n=0}^{N-1} |x[n]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X[k]|^2\,. \tag{29.3.3}$$

e,math,dsp,parseval

s,math,laplace

29.4 Laplace transform (s,math,laplace)

The **Laplace transform** is useful for analyzing the properties of 1D linear time-invariant systems. For signal processing applications, the **bilateral Laplace transform** is usually more relevant than the **unilateral Laplace transform**. Therefore we focus on the bilateral, or **two-sided Laplace transform** here.

Although there exists a formula for the **inverse Laplace transform**, typically all that is needed is to combine Laplace transform properties with known transform pairs such as those summarized below.

s,math,laplace,pair

29.4.1 Laplace transform pairs

signal	transform	ROC	Notes
$h(t)$	$H(s) = \int_{-\infty}^{\infty} h(t)\, \mathrm{e}^{-st}\, \mathrm{d}t$		
$\delta(t)$	1	$\mathbb{C}$	
$\mathrm{e}^{-at}\, \mathrm{step}(t)$	$\frac{1}{s+a}$	$-\mathrm{real}\{a\} < \mathrm{real}\{s\}$	
$\mathrm{e}^{-a\lvert t\rvert}$	$\frac{2a}{a^2 - s^2}$	$\lvert\mathrm{real}\{s\}\rvert < \mathrm{real}\{a\}$	
$\cos(bt)\, \mathrm{e}^{-at}\, \mathrm{step}(t)$	$\frac{s+a}{(s+a)^2 + b^2}$	$-\mathrm{real}\{a\} < \mathrm{real}\{s\}$	Problem 29.2
$\sin(bt)\, \mathrm{e}^{-at}\, \mathrm{step}(t)$	$\frac{b}{(s+a)^2 + b^2}$	$-\mathrm{real}\{a\} < \mathrm{real}\{s\}$	Problem 29.3
$\cos(bt)\, \mathrm{e}^{-a\lvert t\rvert}$	$\frac{2a\left(a^2 + b^2 - s^2\right)}{(a^2 + b^2 - s^2)^2 + (2bs)^2}$	$\lvert\mathrm{real}\{s\}\rvert < \mathrm{real}\{a\}$	Problem 29.4
$\sin(b\lvert t\rvert)\, \mathrm{e}^{-a\lvert t\rvert}$	$\frac{2b\left(a^2 + b^2 + s^2\right)}{(a^2 + b^2 - s^2)^2 + (2bs)^2}$	$\lvert\mathrm{real}\{s\}\rvert < \mathrm{real}\{a\}$	Problem 29.5
$(\cos(at) + \sin(a\lvert t\rvert))\, \mathrm{e}^{-a\lvert t\rvert}$	$\frac{8a^3}{s^4 + 4a^4}$	$\lvert\mathrm{real}\{s\}\rvert < \mathrm{real}\{a\}$	

s,math,laplace,prop

29.4.2 Laplace transform properties

property	signal	transform	ROC
linearity	$ah(t) + bg(t)$	$aH(s) + bG(s)$	$\mathrm{ROC} \supseteq \mathrm{ROC}_1 \cap \mathrm{ROC}_2$
differentiation	$\frac{\mathrm{d}}{\mathrm{d}t} h(t)$	$sH(s)$	$\mathrm{ROC} \supseteq \mathrm{ROC}_h$
convolution	$h(t) * g(t)$	$H(s)G(s)$	$\mathrm{ROC} \supseteq \mathrm{ROC}_h \cap \mathrm{ROC}_g$
time shift	$h(t - \tau)$	$\mathrm{e}^{-\tau s}\, H(s)$	ROC unchanged
time scale	$h(at),\ a \neq 0$	$\frac{1}{\lvert a\rvert} H(s/a)$	$\mathrm{ROC} = \mathrm{ROC}_h / a$
modulation	$\mathrm{e}^{bt}\, h(t)$	$H(s - b)$	$\mathrm{ROC} = \mathrm{ROC}_h + \mathrm{real}\{b\}$
differentiation in s	$-th(t)$	$\frac{\mathrm{d}}{\mathrm{d}s} H(s)$	ROC unchanged
running integration	$\int_{-\infty}^{t} h(t')\, \mathrm{d}t'$	$\frac{1}{s} H(s)$	$\mathrm{ROC} \supseteq \mathrm{ROC}_h \cap \{\mathrm{real}\{s\}\ 0\}$

s,math,diff

29.5 Differential equations (s,math,diff)

If the time-varying state vector $\boldsymbol{x}(t)$ evolves according to the linear, constant-coefficient differential equation

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{u}(t),$$

with initial condition $\boldsymbol{x}(0) = \boldsymbol{x}_0$, then it is readily verified that the solution for $t \geq 0$ is

$$\boldsymbol{x}(t) = \mathrm{e}^{\boldsymbol{A}t}\,\boldsymbol{x}_0 + \int_0^t \mathrm{e}^{\boldsymbol{A}(t-s)}\,\boldsymbol{u}(s)\,\mathrm{d}s\,. \tag{29.5.1}$$

e,math,diff,sol

If $\boldsymbol{A}$ has eigen-decomposition $\boldsymbol{A} = \boldsymbol{V}\,\boldsymbol{\Lambda}\,\boldsymbol{V}^{-1}$, then $\mathrm{e}^{\boldsymbol{A}s} = \boldsymbol{V}\,\boldsymbol{\Lambda}^s\,\boldsymbol{V}^{-1}$. This is useful, for example, when solving the **Bloch equation** in NMR.

If the scalar function $x(t)$ evolves according to the linear, *time-varying* differential equation

$$\dot{x}(t) = a(t)\,x(t) + u(t),$$

with initial condition $x(0) = x_0$, then it is readily verified that the solution for $t \geq 0$ is

$$x(t) = \exp\left(\int_0^t a(s)\,\mathrm{d}s\right) x_0 + \int_0^t \exp\left(\int_s^t a(\tau)\,\mathrm{d}\tau\right) u(s)\,\mathrm{d}s\,.$$

Unfortunately, there is no simple generalization of this result to the vector state case:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}(t)\boldsymbol{x}(t) + \boldsymbol{u}(t)\,.$$

Define $\boldsymbol{B}(t) \triangleq \int_0^t A(s)\,\mathrm{d}s$ and consider the function

$$\boldsymbol{f}(t) = \exp(\boldsymbol{B}(t))\boldsymbol{f}_0 = \left(\boldsymbol{I} + \sum_{k=1}^{\infty} \frac{1}{k!}\left[\boldsymbol{B}(t)\right]^k\right)\boldsymbol{f}_0.$$

Its derivative is

$$\dot{\boldsymbol{f}}(t) = \left(\sum_{k=1}^{\infty} \frac{1}{k!}\left(\frac{\mathrm{d}}{\mathrm{d}t}\left[\boldsymbol{B}(t)\right]^k\right)\right)\boldsymbol{f}_0.$$

This simplifies if the matrix $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{B}(t)$ commutes with $\boldsymbol{B}(t)$, which happens in the scalar case but not in general.

s,math,z

29.6 Z transform (s,math,z)

The **Z transform** is useful for analyzing the properties of discrete-time 1D linear time-invariant systems.

Although there exists a formula for the **inverse Z transform**, typically all that is needed is to combine Z transform properties with known transform pairs such as those summarized below.

s,math,z,pair

29.6.1 Z transform pairs

In the following table, $p = r\,\mathrm{e}^{\imath b} \in \mathbb{C}$, where $r, b \in \mathbb{R}$.

signal	transform	ROC
$h[n]$	$\sum_{n=-\infty}^{\infty} h[n] z^{-n}$	
$\delta[n]$	1	$\mathbb{C}$
$p^n \,\mathrm{step}(n)$	$\frac{1}{1 - pz^{-1}}$	$\lvert p\rvert < \lvert z\rvert$
$p^{\lvert n\rvert}$	$\frac{(p - p^{-1})z^{-1}}{1 - (p + p^{-1})z^{-1} + z^{-2}} = \frac{(p - p^{-1})z^{-1}}{(1 - pz^{-1})(1 - p^{-1}z^{-1})}$	$\lvert p\rvert < \lvert z\rvert < \frac{1}{\lvert p\rvert}$
$\cos(bn)\, r^n \,\mathrm{step}(n)$	$\frac{1 - r\cos(b)\, z^{-1}}{1 - 2r\cos(b)\, z^{-1} + r^2 z^{-2}} = \frac{1 - \mathrm{real}\{p\}\, z^{-1}}{(1 - pz^{-1})(1 - p^* z^{-1})}$	$\lvert r\rvert < \lvert z\rvert$
$\sin(bn)\, r^n \,\mathrm{step}(n)$	$\frac{r\sin(b)\, z^{-1}}{1 - 2r\cos(b)\, z^{-1} + r^2 z^{-2}} = \frac{\mathrm{imag}\{p\}\, z^{-1}}{(1 - pz^{-1})(1 - p^* z^{-1})}$	$\lvert r\rvert < \lvert z\rvert$
$\cos(bn)\, r^{\lvert n\rvert}$	$z^{-1} \frac{(1 + z^{-2})(r - 1/r)\cos(b) - (r^2 - 1/r^2)z^{-1}}{(1 - pz^{-1})(1 - 1/pz^{-1})(1 - p^* z^{-1})(1 - 1/p^* z^{-1})}$	$\lvert r\rvert < \lvert z\rvert < \frac{1}{\lvert r\rvert}$
$\sin(b\,\lvert n\rvert)\, r^{\lvert n\rvert}$	$z^{-1} \frac{(1 + z^{-2})(r + 1/r)\sin(b) - 2\sin(2b)\, z^{-1}}{(1 - pz^{-1})(1 - 1/pz^{-1})(1 - p^* z^{-1})(1 - 1/p^* z^{-1})}$	$\lvert r\rvert < \lvert z\rvert < \frac{1}{\lvert r\rvert}$

s,math,z,prop

29.6.2 Z transform properties

property	signal	transform	ROC
linearity	$ah[n] + bg[n]$	$aH(z) + bG(z)$	$\mathrm{ROC} \supseteq \mathrm{ROC}_1 \cap \mathrm{ROC}_2$
differencing	$h[n] - h[n-1]$	$(1 - z^{-1})H(z)$	$\mathrm{ROC} \supseteq \mathrm{ROC}_h$
convolution	$h[n] * g[n]$	$H(z)G(z)$	$\mathrm{ROC} \supseteq \mathrm{ROC}_h \cap \mathrm{ROC}_g$
time shift	$h[n - m]$	$z^{-m}H(z)$	ROC unchanged, except origin
time reversal	$h[-n]$	$H(z^{-1})$	ROC inverted
modulation	$b^n h[n]$	$H(z/b)$	$\mathrm{ROC} = \mathrm{ROC}_h / \lvert b\rvert$
time expansion (upsampling), $k \in \mathbb{N}$	$h[n/k]\mathbb{I}_{\{n/k \in \mathbb{Z}\}}$	$H(z^k)$	$\mathrm{ROC} = \mathrm{ROC}_h^{1/k}$
differentiation in z	$-nh[n]$	$-\frac{\mathrm{d}}{\mathrm{d}z}H(z)$	ROC unchanged
accumulation	$\sum_{k=-\infty}^{n} h[k]$	$\frac{1}{1 - z^{-1}}H(z)$	$\mathrm{ROC} \supseteq \mathrm{ROC}_h \cap \{\lvert z\rvert > 1\}$

s,math,spline

29.7 B-splines (s,math,spline)

B-spline functions of order $k \geq 0$ with **knots** $t_1, t_2, \ldots, t_{N+k+1}$ are **defined recursively** as follows for $i = 1, \ldots, N$:

$$B_{i,0}(t) = \mathbb{I}_{\{t_i \leq t < t_{i+1}\}}$$
$$B_{i,k}(t) = \frac{t - t_i}{t_{i+k} - t_i} B_{i,k-1}(t) + \frac{t_{i+1+k} - t}{t_{i+1+k} - t_{i+1}} B_{i+1,k-1}(t). \quad (29.7.1)$$

e,math,bspline

In image processing, often the knots are spaced equally.

MIRT *See* `ir_bspline_k.m`.

s,math,gradient

29.8 Gradients (s,math,gradient)

If $f(\boldsymbol{x})$ is a (differentiable) function from $\mathbb{R}^{n_\mathrm{p}}$ to $\mathbb{R}$, then the **row gradient** of f is defined as

$$\nabla f(\boldsymbol{x}) = \left[\frac{\partial}{\partial x_1} f(\boldsymbol{x}), \dots, \frac{\partial}{\partial x_{n_\mathrm{p}}} f(\boldsymbol{x}) \right]. \tag{29.8.1}$$

e,rgrad

The **column gradient**, denoted ∇f is the transpose of the row gradient.

29.8.1 Gradients of linear and quadratic forms

If $f(\boldsymbol{x}) = \mathrm{real}\{\boldsymbol{v}' \boldsymbol{x}\}$ for some vector $\boldsymbol{v} \in \mathbb{C}^{n_\mathrm{p}}$, then (see §29.2):

$$\nabla f(\boldsymbol{x}) = \boldsymbol{v}\,.$$

If $f(\boldsymbol{x}) = \boldsymbol{x}' \boldsymbol{M} \boldsymbol{x}$ for some matrix $\boldsymbol{M} \in \mathbb{R}^{n_\mathrm{p} \times n_\mathrm{p}}$, then

$$\nabla f(\boldsymbol{x}) = (\boldsymbol{M} + \boldsymbol{M}')\boldsymbol{x}. \tag{29.8.2}$$

e,math,gradient,quad

See Appendix 29 for the complex case.

Leibniz's rule:

$$G(x) = \int_{a(x)}^{b(x)} h(x,y)\,\mathrm{d}y \implies \frac{d}{dx} G(x) = h(x,b(x)) \frac{d}{dx} b(x) - h(x,a(x)) \frac{d}{dx} a(x) + \int_{a(x)}^{b(x)} \frac{\partial}{\partial x} h(x,y)\,\mathrm{d}y\,.$$

29.8.2 Taylor series expansions

If $g : \mathbb{R} \to \mathbb{R}$ is an n-times differentiable function, then **Taylor's theorem** is:

$$g(x) = g(a) + \sum_{k=1}^{n-1} \frac{1}{k!} g^{(k)}(a)(x-a)^k + \int_a^x \frac{1}{(n-1)!} (x-t)^{n-1} g^{(n)}(t)\,\mathrm{d}t\,.$$

In particular, for $n = 2$ we have

$$\begin{aligned} g(t) &= g(s) + \dot{g}(s)(t-s) + \int_s^t (t-\tau) \ddot{g}(\tau)\,\mathrm{d}\tau \\ &= g(s) + \dot{g}(s)(t-s) + (t-s)^2 \int_0^1 (1-\alpha) \ddot{g}(\alpha t + (1-\alpha)s)\,\mathrm{d}\alpha\,. \end{aligned}$$

Conversely, if

$$g(t) = g(s) + \dot{g}(s)(t-s) + (t-s)^2 \int_0^1 (1-\alpha)\, \check{c}(\alpha t + (1-\alpha)\, s)\,\mathrm{d}\alpha,$$

then $\ddot{g}(t) = \check{c}(t)$.

If $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable, then the **1st-order Taylor series** expansion of f around a point $\boldsymbol{z}$ is

$$f(\boldsymbol{x}) = f(\boldsymbol{z}) + \left[\int_0^1 \nabla f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{z})\,\mathrm{d}\alpha \right] (\boldsymbol{x} - \boldsymbol{z}). \tag{29.8.3}$$

e,taylor1

If $f : \mathbb{R}^n \to \mathbb{R}$ is twice differentiable, then the **2nd-order Taylor series** expansion of f around a point $\boldsymbol{z}$ is

$$f(\boldsymbol{x}) = f(\boldsymbol{z}) + \nabla f(\boldsymbol{z})(\boldsymbol{x} - \boldsymbol{z}) + (\boldsymbol{x} - \boldsymbol{z})' \left[\int_0^1 (1-\alpha) \nabla^2 f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{z})\,\mathrm{d}\alpha \right] (\boldsymbol{x} - \boldsymbol{z}). \tag{29.8.4}$$

e,taylor2

For functions with complex arguments, see §29.4.

29.8.3 Lipschitz continuity

A function $\boldsymbol{g} : \mathbb{C}^n \to \mathbb{C}^m$ is called **Lipschitz continuous** if there exists a finite real number $\mathcal{L}$ such that

e,math,lipschitz,def

$$\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{z})\|_2 \leq \mathcal{L} \|\boldsymbol{x} - \boldsymbol{z}\|_2, \quad \forall \boldsymbol{x}, \boldsymbol{z} \in \mathbb{C}^n. \tag{29.8.5}$$

Such an $\mathcal{L}$ is called a **Lipschitz constant**. If $\boldsymbol{g}$ is **Lipschitz continuous**, then it is natural to seek the smallest **Lipschitz constant**:

$$\mathcal{L}_* \triangleq \sup_{\boldsymbol{x} \neq \boldsymbol{z}} \frac{\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{z})\|_2}{\|\boldsymbol{x} - \boldsymbol{z}\|_2},$$

which one might call "the" **Lipschitz constant** of $\boldsymbol{g}$.

A function need *not* be differentiable to be Lipschitz continuous. For example, $g(x) = |x|$ is Lipschitz continuous with $\mathcal{L} = 1$. However, any **Lipschitz continuous** function is **absolutely continuous** and thus differentiable **almost everywhere**. We generalize this definition in Definition 29.9.16.

s,math,convex

29.9 Convexity (s,math,convex)

Convex sets and convex functions have important roles in optimization. This section reviews some relevant properties.

s,math,convex,set

29.9.1 Convex sets

d,convex,set

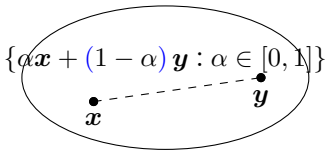
Definition 29.9.1 *A set $\mathcal{C}$ is **convex** iff*

$$\boldsymbol{x}, \boldsymbol{z} \in \mathcal{C} \implies \alpha \boldsymbol{x} + (1-\alpha)\, \boldsymbol{z} \in \mathcal{C}, \qquad \forall \alpha \in [0,1].$$

*In words, for any two points in a convex set, all points on the **line segment** between those points also lie in the set.*

x,convex,ellipse

Example 29.9.2 *An ellipse is a convex set, as illustrated in the following figure.*



x,math,convex,box

Example 29.9.3 *Any **box set** of the following form is convex:*

$$\{\boldsymbol{x} \in \mathbb{R}^n : l_j \le x_j \le u_j,\ j = 1, \dots, n\}\,. \tag{29.9.1}$$

e,math,convex,box

,convex,nonneg,orthant

Example 29.9.4 *The **nonnegative orthant** is the most important convex set used in this book:*

$$\{\boldsymbol{x} \in \mathbb{R}^n : 0 \le x_j < \infty,\ j = 1, \dots, n\}\,. \tag{29.9.2}$$

e,math,convex,nonneg,orthant

math,convex,projection

29.9.2 Convex projections

A particularly important property of convex sets is that for any point $\boldsymbol{x}$ outside the set, there is a unique point $\boldsymbol{x}^{(\star)}$ within the set that is closest to $\boldsymbol{x}$, as established by the following theorem from functional analysis [1, p. 69].

t,math,convex,closest

Theorem 29.9.5 *Let $\mathcal{C}$ be a nonempty closed convex subset of a Hilbert space H (such as $\mathbb{R}^n$ or $\mathbb{C}^n$) with associated inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$. For any $\boldsymbol{x} \in H$, there is a unique vector $\boldsymbol{x}^{(\star)} \in \mathcal{C}$ such that*

$$\|\boldsymbol{x} - \boldsymbol{x}^{(\star)}\| \le \|\boldsymbol{x} - \boldsymbol{z}\|\,, \qquad \forall \boldsymbol{z} \in \mathcal{C}.$$

Furthermore, $\boldsymbol{x}^{(\star)}$ is characterized (in a necessary and sufficient sense) by

$$\operatorname{real}\{\langle \boldsymbol{x} - \boldsymbol{x}^{(\star)},\ \boldsymbol{z} - \boldsymbol{x}^{(\star)} \rangle\} \le 0, \qquad \forall \boldsymbol{z} \in \mathcal{C}. \tag{29.9.3}$$

e,math,convex,closest,inprod

This theorem is a close relative of the **projection theorem**. Because of its existence and uniqueness results, it is valid to define the following **projector** onto a convex set $\mathcal{C}$:

$$\mathcal{P}_{\mathcal{C}}(\boldsymbol{x}) \triangleq \arg\min_{\boldsymbol{z} \in \mathcal{C}} \|\boldsymbol{z} - \boldsymbol{x}\|\,. \tag{29.9.4}$$

e,math,convex,projector

This function gives the closest point in $\mathcal{C}$ to $\boldsymbol{x}$.

h,convex,projector,box

Example 29.9.6 *If $\mathcal{C}$ is the box set defined in (29.9.1) and $\boldsymbol{x}^{(\star)} = \mathcal{P}_{\mathcal{C}}(\boldsymbol{x})$, then*

$$x_j^\star = \begin{cases} l_j, & x_j < l_j \\ x_j, & l_j \le x_j \le u_j \\ u_j, & x_j > u_j. \end{cases}$$

s,math,convex,fun

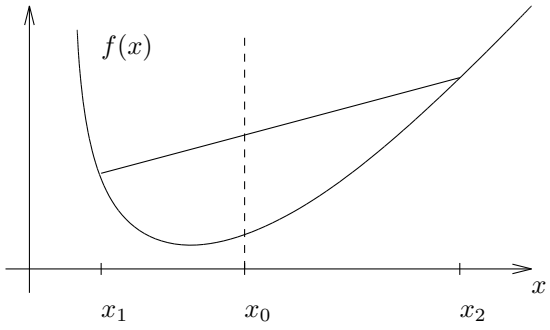
29.9.3 Convex functions (s,math,convex,fun)

d,convex,fun

Definition 29.9.7 *For a convex set $\mathcal{D} \subset \mathbb{R}^n$, a real functional $f : \mathcal{D} \to \mathbb{R}$ is called a* ***convex function*** *on $\mathcal{D}$ iff*

$$f(\alpha \boldsymbol{x} + (1-\alpha)\,\boldsymbol{z}) \le \alpha\, f(\boldsymbol{x}) + (1-\alpha)\, f(\boldsymbol{z}), \qquad \forall \alpha \in [0,1], \qquad \forall \boldsymbol{x}, \boldsymbol{z} \in \mathcal{D}.$$

The function f is called **strictly convex** iff the inequality is strict for all $\boldsymbol{x} \neq \boldsymbol{z}$ and $\alpha \in (0,1)$.



fig,convex

Figure 29.9.1: Illustration of a 1D convex function.

An alternative (equivalent) definition is that a function is convex iff its **epigraph** (the set of points on or above the graph of the function, *i.e.*, $\{(\boldsymbol{x}, t) : \boldsymbol{x} \in \mathcal{D}, f(\boldsymbol{x}) \le t \in \mathbb{R}\}$) is a **convex set** [2, p. 75].

d,concave,fun

Definition 29.9.8 *A function f is called* ***concave*** *if its negative, $-f$, is convex.*
A ***strictly concave*** *function is defined by analogy.*

x,math,convex,parabola

Example 29.9.9 *A parabola $f(x) = a + bx + cx^2$ with $c \ge 0$ is convex on $\mathbb{R}$, because for $\alpha \in [0,1]$:*

$$\alpha f(x) + (1-\alpha)\, f(y) - f(\alpha x + (1-\alpha)\, y) = c\alpha\,(1-\alpha)\,(x-y)^2 \ge 0.$$

x,math,convex,power

Example 29.9.10 *The function $f(x) = |x|^p$ is convex for $p \ge 1$ [2, p. 71] and is strictly convex for $p > 1$. (See Problem 29.7.)*

For generalizations of convexity, see [3] (**g-convex**) and [2] (**quasi-convex**).

29.9.3.1 Minimizers of convex functions

d,math,local,min

Definition 29.9.11
$\boldsymbol{x}^{(\star)}$ is a ***local minimizer*** *of $f(\boldsymbol{x})$ with respect to norm $\|\cdot\|$ iff $\exists \varepsilon > 0$ s.t. $\|\boldsymbol{x} - \boldsymbol{x}^{(\star)}\| \le \varepsilon \implies f(\boldsymbol{x}^{(\star)}) \le f(\boldsymbol{x})$ [4, p. 19].*
$\boldsymbol{x}^{(\star)} \in \mathcal{D}$ is a ***global minimizer*** *of $f(\boldsymbol{x})$ over the domain $\mathcal{D}$ iff $f(\boldsymbol{x}^{(\star)}) \le f(\boldsymbol{x})\ \forall \boldsymbol{x} \in \mathcal{D}$.*

Convex functions are particularly important for optimization problems because of the following properties.

- Any **local minimizer** of a convex function is a global minimizer.
- A global minimizer of a strictly convex function is **unique**.

The converse of uniqueness is not true; *i.e.*, there are many functions that have unique global minimizers that are not strictly convex, such as the ℓ_1 norm: $f(\boldsymbol{x}) = \|\boldsymbol{x}\|_1$, and even non-convex, such $f(x) = \sqrt{|x|}$. Convexity is sufficient *but not necessary* for ensuring uniqueness of a global minimizer.

Neither of these properties ensures that local or global minimizers **exist** for convex functions. For example, the 1D function $f(x) = \mathrm{e}^x$ is strictly convex but has no local (or global) minimizers on $\mathbb{R}$.

For differentiable convex functions on $\mathbb{R}^n$ we can characterize the **global minimizer**(s) using the gradient.

- If $\Psi(\boldsymbol{x})$ is **convex** on $\mathbb{R}^n$, then $\nabla \Psi(\boldsymbol{x}) = \boldsymbol{0}$ iff $\boldsymbol{x}$ is a global minimizer of Ψ.
- If $\Psi(\boldsymbol{x})$ is **strictly convex** on $\mathbb{R}^n$, then $\nabla \Psi(\boldsymbol{x}) = \boldsymbol{0}$ iff $\boldsymbol{x}$ is the (unique) global minimizer of Ψ.

In other words, to find an **unconstrained minimizer** of a convex function over $\mathbb{R}^n$, it suffices to find a value of $\boldsymbol{x}$ where the gradient is zero. In contrast, for general (possibly non-convex) functions, finding a point $\boldsymbol{x}$ where the gradient is zero tells us very little because $\boldsymbol{x}$ could be a local minimizer or a local maximizer or neither (a saddle point).

29.9.3.2 Assessing convexity using properties

One way to determine if a function is convex is to resort to the definition. Often it is simpler to use *properties* of the function such as the following.

- A function $f(\boldsymbol{x})$ is convex on $\mathcal{D}$ iff its restriction $g(t) = f(\boldsymbol{x} + t\boldsymbol{z})$ to a line in $\mathcal{D}$ is convex for any $\boldsymbol{x}$ and $\boldsymbol{z}$ in $\mathcal{D}$ [2, p. 68].
- If f is twice differentiable, then f is convex if and only if its **Hessian matrix** is positive-semidefinite:

$$\nabla^2 f \succeq \mathbf{0}. \tag{29.9.5}$$

e,math,convex,fun,hess,mgeq

(See Lemma 29.5.7 and Lemma 29.5.8.)

- If f is twice differentiable, then f is **strictly convex** if its Hessian matrix is positive definite:

$$\nabla^2 f \succ \mathbf{0}. \tag{29.9.6}$$

e,math,convex,fun,hess,mgt

However, the converse is not true. Example: $f(x) = x^4$ is strictly convex, but its Hessian (second derivative) is not positive everywhere.

Often convex functions are constructed from other (convex) functions and the following properties apply.

- If f is convex on $\mathbb{R}^n$, then $g(\boldsymbol{x}) \triangleq f(\boldsymbol{M}\boldsymbol{x}+\boldsymbol{b})$ is convex on $\mathbb{R}^m$ for any matrix $\boldsymbol{M} \in \mathbb{R}^{n \times m}$ and vector $\boldsymbol{b} \in \mathbb{R}^n$. In other words, convexity is preserved under affine transformations. (See Problem 29.6.)
- If $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ are convex, then so is their point-wise maximum: $h(\boldsymbol{x}) = \max\{f(\boldsymbol{x}), g(\boldsymbol{x})\}$.
- Sums of convex functions are convex.
- Sums of strictly convex functions are strictly convex.
- The sum of a convex function and a strictly convex function is strictly convex.

x,convex,fun,sum

Example 29.9.12 *The following cost function is **strictly convex** for any $\beta > 0$: $\Psi(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \beta \|\boldsymbol{x}\|_2^2$. The Hessian of the Tikhonov regularizer $g(\boldsymbol{x}) = \beta \|\boldsymbol{x}\|_2^2$ is $\nabla^2 g = \beta \boldsymbol{I}$, which is positive definite, so g is strictly convex. The data-fit term $\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2$ is convex by the affine property above, so the sum is strictly convex.*

29.9.3.3 Properties of convex functions

- By induction, one can show that if f is convex and $\sum_{k=1}^K \alpha_k = 1$ for $\alpha_k \geq 0$, then f obeys the following **convexity inequality**

$$f\left(\sum_{k=1}^K \alpha_k \boldsymbol{x}_k\right) \leq \sum_{k=1}^K \alpha_k f(\boldsymbol{x}_k), \tag{29.9.7}$$

e,math,convex,convexity,inequality

for $\boldsymbol{x}_k$ values in $\mathcal{D}$, the domain of f.

- A special case of the preceding result is **Jensen's inequality**. If $f : \mathcal{D} \to \mathbb{R}$ is a convex function and $\boldsymbol{X} \in \mathbb{R}^n$ is a random vector with $\mathsf{P}\{\boldsymbol{X} \in \mathcal{D}\} = 1$, then

$$f(\mathsf{E}[\boldsymbol{X}]) \leq \mathsf{E}[f(\boldsymbol{X})]. \tag{29.9.8}$$

e,jensen

- The **sublevel sets** $\{\boldsymbol{x} \in \mathcal{D} : f(\boldsymbol{x}) < a\}$ of a convex function f are convex sets for any $a \in \mathbb{R}$.

x,convex,fun,mean

Example 29.9.13 *Applying (29.9.7) to the function $f(x) = -\log(x)$ yields the weighted **arithmetic-geometric mean inequality** [5, p. 53a]:*

$$\sum_k \alpha_k x_k \geq \prod_k x_k^{\alpha_k}, \qquad \forall x_k \geq 0,$$

if $\sum_k \alpha_k = 1$ and $\alpha_k \geq 0$, with equality iff all x_k values are equal.

29.9.3.4 Properties of differentiable convex functions

- If f is convex on $\mathcal{D}$ and differentiable at $\boldsymbol{z} \in \mathcal{D}$, then it satisfies the following **support property**:

e,math,convex,above,tangent
$$f(\boldsymbol{x}) \geq f(\boldsymbol{z}) + \langle \nabla f(\boldsymbol{z}),\, \boldsymbol{x} - \boldsymbol{z} \rangle, \qquad \forall \boldsymbol{x} \in \mathcal{D}. \tag{29.9.9}$$

  In other words, a convex function lies above the tangent plane at any point.

- Conversely, if f is differentiable on $\mathcal{D}$ and satisfies (29.9.9) everywhere, then f is convex [2, p. 70].

The following lemmas are used to prove convergence of the **gradient projection** algorithm in §12.2.

l,math,convex,polyak,2
Lemma 29.9.14 *[6, p. 10] If $f : \mathcal{D} \to \mathbb{R}$ is convex and differentiable on $\mathcal{D}$, then*

$$\langle \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{z}),\, \boldsymbol{x} - \boldsymbol{z} \rangle \geq 0, \qquad \forall \boldsymbol{x}, \boldsymbol{z} \in \mathcal{D}.$$

Proof:
f is convex $\implies f(\boldsymbol{z}) \geq f(\boldsymbol{x}) + \langle \nabla f(\boldsymbol{x}),\, \boldsymbol{z} - \boldsymbol{x} \rangle$ and $f(\boldsymbol{x}) \geq f(\boldsymbol{z}) + \langle \nabla f(\boldsymbol{z}),\, \boldsymbol{x} - \boldsymbol{z} \rangle$. Now add. □

d,math,convex,lips,mn
Definition 29.9.15 *A function $\boldsymbol{g} : \mathbb{C}^n \to \mathbb{C}^m$ is Lipschitz continuous w.r.t. norms $\|\cdot\|_m$ and $\|\cdot\|_n$ on $\mathbb{C}^m$ and $\mathbb{C}^n$ respectively iff*

e,math,convex,fun,lip,mn
$$\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{z})\|_m \leq \|\boldsymbol{x} - \boldsymbol{z}\|_n, \quad \forall \boldsymbol{x}, \boldsymbol{z} \in \mathbb{C}^n. \tag{29.9.10}$$

d,math,convex,ab,lips
Definition 29.9.16 *A function $\boldsymbol{g} : \mathbb{C}^n \to \mathbb{C}^m$ is $(\boldsymbol{A}, \boldsymbol{B})$-Lipschitz continuous iff*

e,math,convex,fun,lip,ab
$$\|\boldsymbol{A}\,(\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{z}))\|_2 \leq \|\boldsymbol{B}\,(\boldsymbol{x} - \boldsymbol{z})\|_2, \quad \forall \boldsymbol{x}, \boldsymbol{z} \in \mathbb{C}^n, \tag{29.9.11}$$

where $\boldsymbol{A}$ has m columns and $\boldsymbol{B} \in \mathbb{C}^{n \times n}$ is invertible.

Our primary interest in Lipschitz functions will be gradients of cost functions, which are mappings from $\mathbb{C}^{n_\mathrm{p}}$ into $\mathbb{C}^{n_\mathrm{p}}$ (or $\mathbb{R}^{n_\mathrm{p}}$ into $\mathbb{R}^{n_\mathrm{p}}$). The following definition generalizes (29.8.5) and is a special case of Definition 29.9.16.

d,math,convex,s,lips
Definition 29.9.17 *A function $\boldsymbol{g} : \mathbb{C}^{n_\mathrm{p}} \to \mathbb{C}^{n_\mathrm{p}}$ is $\boldsymbol{S}$-Lipschitz continuous, for an invertible $n_\mathrm{p} \times n_\mathrm{p}$ matrix $\boldsymbol{S}$, iff*

e,math,convex,lip,s
$$\left\|\boldsymbol{S}^{-1}\,(\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{z}))\right\|_2 \leq \|\boldsymbol{S}'\,(\boldsymbol{x} - \boldsymbol{z})\|_2, \quad \forall \boldsymbol{x}, \boldsymbol{z} \in \mathbb{C}^{n_\mathrm{p}}. \tag{29.9.12}$$

This definition is a strict generalization of the usual definition of Lipschitz continuity because if (29.9.12) holds, then

$$\begin{aligned} \|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{z})\|_2 &= \left\|\boldsymbol{S}\boldsymbol{S}^{-1}\,(\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{z}))\right\|_2 \leq \|\boldsymbol{S}\|_2 \left\|\boldsymbol{S}^{-1}\,(\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}(\boldsymbol{z}))\right\|_2 \leq \|\boldsymbol{S}\|_2 \|\boldsymbol{S}'\,(\boldsymbol{x} - \boldsymbol{z})\|_2 \\ &\leq \|\boldsymbol{S}\|_2 \|\boldsymbol{S}'\|_2 \|\boldsymbol{x} - \boldsymbol{z}\|_2, \end{aligned}$$

so $\boldsymbol{g}(\cdot)$ is Lipschitz continuous with Lipschitz constant $L = \|\boldsymbol{S}\|_2 \|\boldsymbol{S}'\|_2$.

t,math,convex,lip,hess
Theorem 29.9.18 *If Ψ is twice differentiable, then the Hessian bound $\|\boldsymbol{S}^{-1}\,\nabla^2 \Psi(\boldsymbol{x})\,\boldsymbol{S}^{-\mathrm{H}}\|_2 \leq 1,\ \forall \boldsymbol{x} \in \mathbb{R}^{n_\mathrm{p}}$ holds iff its gradient is $\boldsymbol{S}$-Lipschitz continuous per (29.9.12).*
Proof (extended from [7, p. 21]):
Using Taylor expansion with remainder (29.8.4) and the triangle inequality shows (29.9.12) as follows:

$$\begin{aligned} \left\|\boldsymbol{S}^{-1}\,(\nabla \Psi(\boldsymbol{z}) - \nabla \Psi(\boldsymbol{x}))\right\| &= \left\|\boldsymbol{S}^{-1} \int_0^1 \nabla^2 \Psi(\boldsymbol{x} + \tau\,(\boldsymbol{z} - \boldsymbol{x}))\,\boldsymbol{S}^{-\mathrm{H}} \boldsymbol{S}'\,(\boldsymbol{z} - \boldsymbol{x})\,\mathrm{d}\tau\right\| \\ &\leq \left(\int_0^1 \|\boldsymbol{S}^{-1}\,\nabla^2 \Psi(\boldsymbol{x} + \tau\,(\boldsymbol{z} - \boldsymbol{x}))\,\boldsymbol{S}^{-\mathrm{H}}\|_2\,\mathrm{d}\tau\right) \|\boldsymbol{S}'\,(\boldsymbol{z} - \boldsymbol{x})\| \\ &\leq \|\boldsymbol{S}'\,(\boldsymbol{z} - \boldsymbol{x})\|. \end{aligned}$$

Lemma 29.9.19 provides a (generalized) converse. □

ath,convex,gp,mnorm,ab
Lemma 29.9.19 *If $\boldsymbol{g} : \mathbb{C}^n \to \mathbb{C}^m$ is differentiable at $\boldsymbol{x} \in \mathbb{C}^n$ and $(\boldsymbol{A}, \boldsymbol{B})$-Lipschitz continuous per (29.9.11), then*

$$\|\boldsymbol{A} \nabla \boldsymbol{g}(\boldsymbol{x})\,\boldsymbol{B}^{-1}\| \leq 1.$$

Proof:
Because $\boldsymbol{g}$ is differentiable at $\boldsymbol{x}$, there exists $\boldsymbol{H}(\boldsymbol{x}) = \nabla \boldsymbol{g}(\boldsymbol{x}) \in \mathbb{C}^{m \times n}$ such that

$$\forall \varepsilon > 0,\ \exists \delta_\varepsilon > 0 \text{ s.t. } \|\boldsymbol{d}\| \leq \delta_\varepsilon \implies \|\boldsymbol{g}(\boldsymbol{x}+\boldsymbol{d}) - \boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{H}(\boldsymbol{x})\boldsymbol{d}\| \leq \varepsilon \|\boldsymbol{d}\| .$$

Suppose $\|\boldsymbol{A}\boldsymbol{H}\boldsymbol{B}^{-1}\| > 1$. Then there exists $\boldsymbol{z} \neq \boldsymbol{0}$ and $c > 0$ such that $\left\|\boldsymbol{A}\boldsymbol{H}\boldsymbol{B}^{-1}\boldsymbol{z}\right\| = (1+c)\|\boldsymbol{z}\|$. Consider $\varepsilon = c/(2\|\boldsymbol{A}\|\|\boldsymbol{B}^{-1}\|)$ and choose α such that $\alpha \left\|\boldsymbol{B}^{-1}\boldsymbol{z}\right\| < \delta_\varepsilon$. (The case where $\|\boldsymbol{A}\| = 0$ is trivial.) Defining $\boldsymbol{d} = \alpha \boldsymbol{B}^{-1}\boldsymbol{z}$ and using the triangle inequality:

$$\begin{aligned}(1+c)\|\alpha \boldsymbol{z}\| = \left\|\boldsymbol{A}\boldsymbol{H}\boldsymbol{B}^{-1}\alpha\boldsymbol{z}\right\| &\leq \|\boldsymbol{A}\left(\boldsymbol{g}(\boldsymbol{x}+\boldsymbol{d}) - \boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{H}\boldsymbol{d}\right)\| + \|\boldsymbol{A}\left(\boldsymbol{g}(\boldsymbol{x}+\boldsymbol{d}) - \boldsymbol{g}(\boldsymbol{x})\right)\| \\ &\leq \|\boldsymbol{A}\|\varepsilon\|\boldsymbol{d}\| + \|\boldsymbol{B}\boldsymbol{d}\| \leq (c/2)\|\alpha\boldsymbol{z}\| + \|\alpha\boldsymbol{z}\|,\end{aligned}$$

a contradiction. So we must have $\|\boldsymbol{A}\boldsymbol{H}\boldsymbol{B}^{-1}\| \leq 1$. □

th,convex,lip,majorize

Corollary 29.9.20 *If Ψ is twice differentiable and convex, and $\boldsymbol{S}$ is invertible, then the Lipschitz condition (29.9.12) holds iff $\nabla^2\Psi(\boldsymbol{x}) \preceq \boldsymbol{S}\boldsymbol{S}'$, $\forall \boldsymbol{x}$. (Problem 29.9.)*

In other words, for a twice differentiable, convex cost function Ψ, the Lipschitz condition (29.9.12) is equivalent to a bound on the curvature (Hessian) of Ψ.

The condition in Definition 29.9.17 leads to the following inequality related to Lemma 29.9.14. (See Problem 29.8.)

math,convex,gp,upper,s

Lemma 29.9.21 *If Ψ is convex and differentiable and its gradient is $\boldsymbol{S}$-Lipschitz continuous per Definition 29.9.17, then*

$$\langle \nabla\Psi(\boldsymbol{x}) - \nabla\Psi(\boldsymbol{z}),\ \boldsymbol{x} - \boldsymbol{z}\rangle \leq \|\boldsymbol{S}'(\boldsymbol{x}-\boldsymbol{z})\|^2 . \tag{29.9.13}$$

e,math,convex,gp,upper,s

math,convex,gp,lower,s

Lemma 29.9.22 *If Ψ is convex and differentiable and its gradient is $\boldsymbol{S}$-Lipschitz continuous per Definition 29.9.17, then we can strengthen Lemma 29.9.14 to*

$$\langle \nabla\Psi(\boldsymbol{x}) - \nabla\Psi(\boldsymbol{z}),\ \boldsymbol{x} - \boldsymbol{z}\rangle \geq \left\|\boldsymbol{S}^{-1}\left(\nabla\Psi(\boldsymbol{x}) - \nabla\Psi(\boldsymbol{z})\right)\right\|^2 . \tag{29.9.14}$$

e,math,convex,gp,lower

Proof (extended from [6, p. 24]):
Consider first the case where Ψ is twice differentiable. Then by the 2nd-order Taylor expansion (29.4.3): $\nabla\Psi(\boldsymbol{x}) = \nabla\Psi(\boldsymbol{z}) + \boldsymbol{H}(\boldsymbol{x},\boldsymbol{z})(\boldsymbol{x}-\boldsymbol{z})$ where $\boldsymbol{H}(\boldsymbol{x},\boldsymbol{z}) = \int_0^1 \nabla^2\Psi(\tau\boldsymbol{x} + (1-\tau)\boldsymbol{z})\,\mathrm{d}\tau \succeq \boldsymbol{0}$ by the convexity of Ψ. Defining $\boldsymbol{A} \triangleq \boldsymbol{A}(\boldsymbol{x},\boldsymbol{z}) = \boldsymbol{S}^{-1}\boldsymbol{H}(\boldsymbol{x},\boldsymbol{z})\boldsymbol{S}^{-\mathrm{H}} \succeq \boldsymbol{0}$, by Lemma 29.9.19 we have $\|\boldsymbol{A}\| \leq 1$. Note that $\boldsymbol{A}^2 \preceq \|\boldsymbol{A}\|\boldsymbol{A} \preceq \boldsymbol{A}$, so $\boldsymbol{w}'\boldsymbol{A}\boldsymbol{w} \geq \|\boldsymbol{A}\boldsymbol{w}\|^2$. Thus with $\boldsymbol{w} = \boldsymbol{S}'(\boldsymbol{x}-\boldsymbol{z})$:

$$\begin{aligned}\langle \nabla\Psi(\boldsymbol{x}) - \nabla\Psi(\boldsymbol{z}),\ \boldsymbol{x}-\boldsymbol{z}\rangle &= \langle \boldsymbol{H}(\boldsymbol{x}-\boldsymbol{z}),\ \boldsymbol{x}-\boldsymbol{z}\rangle = \left\langle \boldsymbol{S}^{-1}\boldsymbol{H}\boldsymbol{S}^{-\mathrm{H}}\boldsymbol{w},\ \boldsymbol{w}\right\rangle = \boldsymbol{w}'\boldsymbol{A}\boldsymbol{w} \\ &\geq \|\boldsymbol{A}\boldsymbol{w}\|^2 = \left\|\boldsymbol{S}^{-1}\boldsymbol{H}(\boldsymbol{x}-\boldsymbol{z})\right\|^2 = \left\|\boldsymbol{S}^{-1}\left(\nabla\Psi(\boldsymbol{x}) - \nabla\Psi(\boldsymbol{z})\right)\right\|^2 .\end{aligned}$$

If Ψ is not twice differentiable[1], then consider the smoothed function

$$\Psi_\varepsilon(\boldsymbol{x}) = \int \Psi(\boldsymbol{x}-\boldsymbol{z})\, p(\boldsymbol{z}/\varepsilon)\varepsilon^{-n_\mathrm{p}}\,\mathrm{d}\boldsymbol{z}$$

where $p(\cdot)$ is a smooth (*i.e.*, C^∞) kernel with finite support and unit integral. Clearly Ψ_ε is convex and twice differentiable and its gradient satisfies the same Lipschitz condition as $\nabla\Psi$. Thus Ψ_ε satisfies the inequality (29.9.14), and taking the limit as $\varepsilon \to 0^+$ establishes (29.9.14) for Ψ. □

th,convex,fun,lower2,s

Lemma 29.9.23 *If Ψ is convex and differentiable and its gradient is $\boldsymbol{S}$-Lipschitz continuous per Definition 29.9.17, then [7, Thm. 2.1.5]:*

$$\Psi(\boldsymbol{x}) \leq \Psi(\boldsymbol{z}) + \langle \nabla\Psi(\boldsymbol{z}),\ \boldsymbol{x}-\boldsymbol{z}\rangle + \frac{1}{2}\left\|\boldsymbol{S}\left(\nabla\Psi(\boldsymbol{x}) - \nabla\Psi(\boldsymbol{z})\right)\right\|^2 .$$

This property was used as a definition in [8].

[1] Thanks to Arkadi Nemirovski for help with this proof, in a 2001-08-20 email to Matt Jacobson.

s,math,convex,conj

29.9.3.5 Convex conjugate

The **convex conjugate** of a function, also known as the **Legendre-Fenchel transformation**, is used for deriving some optimization methods. (See §14.8.) The definition applies to general normed spaces but here we consider only real-valued functions on $\mathbb{R}^n$. For such functions, the **convex conjugate** is

e,math,convex,conj

$$f^\star(\boldsymbol{z}) \triangleq \sup_{\boldsymbol{x}} \left(\langle \boldsymbol{z}, \boldsymbol{x} \rangle - f(\boldsymbol{x}) \right), \quad \forall \boldsymbol{z} \in \mathbb{R}^n. \tag{29.9.15}$$

x,math,convex,conj,abs

Example 29.9.24 *The convex conjugate of the absolute value function $f(x) = |x|$ is*

$$f^\star(z) = \begin{cases} 0, & |z| \le 1 \\ \infty, & |z| > 1. \end{cases}$$

math,convex,conj,power

Example 29.9.25 *The convex conjugate of a power function*

$$f(x) = \frac{1}{p} |x|^p, \quad 1 < p < \infty$$

is

$$f^\star(z) = \frac{1}{q} |z|^q, \quad \frac{1}{p} + \frac{1}{q} = 1.$$

Properties of convex conjugates include the following.

- $f^\star(\boldsymbol{x})$ is convex on $\mathbb{R}^n$.
- If $f(\boldsymbol{x})$ is a **proper**, **convex**, and **lower-semicontinuous** function, then $f^{\star\star} = f$ by the **Fenchel-Moreau theorem**, so

e,math,convex,conj,min

$$f(\boldsymbol{x}) = \sup_{\boldsymbol{z}} \left(\langle \boldsymbol{z}, \boldsymbol{x} \rangle - f^\star(\boldsymbol{z}) \right). \tag{29.9.16}$$

In particular we have the following inequality:

e,math,convex,conj,geq

$$f(\boldsymbol{x}) \ge \langle \boldsymbol{z}, \boldsymbol{x} \rangle - f^\star(\boldsymbol{z}), \quad \forall \boldsymbol{z} \in \mathbb{R}^n. \tag{29.9.17}$$

s,math,convex,prox

29.9.3.6 Proximal mapping

If f is a convex function, then the **proximal operator** or **proximal mapping Moreau proximity operator** is defined [9–12] as

e,math,convex,fun,prox

$$\operatorname{prox}_f(\boldsymbol{z}) \triangleq \arg\min_{\boldsymbol{x}} \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{z}\|_2^2 + f(\boldsymbol{x}). \tag{29.9.18}$$

Because $\|\boldsymbol{x} - \boldsymbol{z}\|_2$ is strictly convex and $f(\boldsymbol{x})$ is convex, their sum is strictly convex so that sum has a unique minimizer for any $\boldsymbol{z}$. This operator is particularly useful for the **proximal gradient method** (**PGM**) for non-smooth optimization and its relatives [13, 14].

If f is differentiable, then by differentiating (29.9.18) we have

$$\mathbf{0} = \boldsymbol{x} - \boldsymbol{z} + \nabla f(\boldsymbol{x}) \Big|_{\boldsymbol{x} = \operatorname{prox}_f(\boldsymbol{z})} \implies \operatorname{prox}_f(\boldsymbol{z}) = (\boldsymbol{I} + \nabla f)^{-1}(\boldsymbol{z}).$$

This relationship holds even for nonsmooth convex functions using the **subdifferential**, so

$$\operatorname{prox}_f = (\mathrm{Id} + \partial f)^{-1},$$

where Id denotes the identity operator (not the identity matrix) and the inverse above denotes a function inverse, not a matrix inverse.

convex,moreau,envelope

29.9.3.7 Moreau envelope

For a convex function f and $\lambda > 0$, the **Moreau envelope** of f with parameter λ is defined [9–12, 15] [16, Sect. 12.4] by

e,math,convex,fun,mor,env

$$f_\lambda(\boldsymbol{z}) \triangleq \inf_{\boldsymbol{x}} f(\boldsymbol{x}) + \frac{1}{2\lambda} \|\boldsymbol{x} - \boldsymbol{z}\|_2^2. \tag{29.9.19}$$

Comparing to the proximity operator (29.9.18), we see that

$$f_\lambda(\boldsymbol{z}) = f\big(\operatorname{prox}_{\lambda f}(\boldsymbol{z})\big) + \frac{1}{2\lambda} \left\| \operatorname{prox}_{\lambda f}(\boldsymbol{z}) - \boldsymbol{z} \right\|_2^2.$$

The **Moreau envelope** is also the **infimal convolution** of $f(\cdot)$ with $\frac{1}{2\lambda} \|\cdot\|_2^2$.

s,math,convex,min

29.9.3.8 Minimization of convex functions over convex sets

Many of the image reconstruction problems described in this *book* require finding the minimizer $\hat{\boldsymbol{x}}$ of a convex cost function $\Psi(\boldsymbol{x})$ over a convex set $\mathcal{C} \subset \mathbb{R}^{n_\mathrm{p}}$, *i.e.*,

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x} \in \mathcal{C}} \Psi(\boldsymbol{x}) \,.$$

In general the minimizer of $\Psi(\boldsymbol{x})$ may not be unique even over $\mathcal{C}$, so we define the solution set

$$\mathcal{X}^{(\star)} \triangleq \{\boldsymbol{x}^{(\star)} \in \mathcal{C} : \Psi(\boldsymbol{x}^{(\star)}) \le \Psi(\boldsymbol{x}), \ \forall \boldsymbol{x} \in \mathcal{C}\} \,.$$

Constrained minimization problems require more work than simply equating the gradient to zero. When Ψ is a convex cost function and $\mathcal{C}$ is a convex set, the constrained minimizers are characterized by the following result.

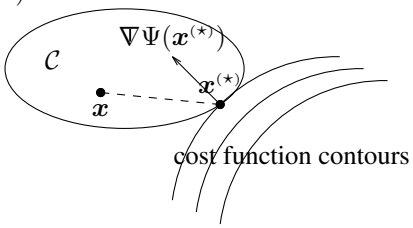
t,math,convex,min

Theorem 29.9.26 *[6, p. 203] If Ψ is a differentiable convex function and $\mathcal{C}$ is a convex set, then*

$$\boldsymbol{x}^{(\star)} \in \mathcal{X}^{(\star)} \iff \nabla \Psi(\boldsymbol{x}^{(\star)}) \, (\boldsymbol{x} - \boldsymbol{x}^{(\star)}) \ge 0, \qquad \forall \boldsymbol{x} \in \mathcal{C}.$$

Proof of $\Leftarrow$ [6, p. 200]:
Suppose $\nabla \Psi(\boldsymbol{x}^{(\star)}) \, (\boldsymbol{x}^{(0)} - \boldsymbol{x}^{(\star)}) < 0$ for some $\boldsymbol{x}^{(0)} \in \mathcal{C}$. Define $\boldsymbol{x}_\alpha = (1-\alpha)\, \boldsymbol{x}^{(\star)} + \alpha \boldsymbol{x}^{(0)} = \boldsymbol{x}^{(\star)} + \alpha(\boldsymbol{x}^{(0)} - \boldsymbol{x}^{(\star)})$ and note $\boldsymbol{x}_\alpha \in \mathcal{C}$ because $\mathcal{C}$ is convex. Then by Taylor series: $\Psi(\boldsymbol{x}_\alpha) = \Psi(\boldsymbol{x}^{(\star)}) + \alpha \nabla \Psi(\boldsymbol{x}^{(\star)}) \, (\boldsymbol{x}_\alpha - \boldsymbol{x}^{(\star)}) + o(\alpha)$ so $\Psi(\boldsymbol{x}_\alpha) < \Psi(\boldsymbol{x}^{(\star)})$ for α sufficiently small, a contradiction. □

The preceding characterization is easily understood geometrically: for $\boldsymbol{x} \in \mathcal{C}$, the vector $\boldsymbol{x} - \boldsymbol{x}^{(\star)}$ has an angle less than $\pi/2$ away from the gradient $\nabla \Psi(\boldsymbol{x}^{(\star)})$.



s,math,convex,min,max

29.9.3.9 Exchanging order of minimization and maximization

In some optimization problems one would like to exchange the order of minimzation and maximization as follows:

$$\min_{\boldsymbol{x} \in \mathcal{X}} \sup_{\boldsymbol{y} \in \mathcal{Y}} \Psi(\boldsymbol{x}, \boldsymbol{y}) \stackrel{?}{=} \sup_{\boldsymbol{y} \in \mathcal{Y}} \min_{\boldsymbol{x} \in \mathcal{X}} \Psi(\boldsymbol{x}, \boldsymbol{y}) \,.$$

Sufficient conditions (including convexity in $\boldsymbol{x}$ and concavity in $\boldsymbol{y}$) are given in [17, 18].

29.10 Minimizers with nonnegativity constraints (s,math,min,nonneg)

s,math,min,nonneg
t,math,min,nonneg

Theorem 29.10.1 *Let $f(\boldsymbol{x})$ be a differentiable function defined over the **nonnegative orthant** $\mathbb{R}_+^{n_\mathrm{p}}$. Let $\boldsymbol{D} = \mathsf{diag}\{d_j\}$ denote a diagonal $n_\mathrm{p} \times n_\mathrm{p}$ matrix with nonnegative diagonal elements. Let $\boldsymbol{x}^{(\star)} \in \mathbb{R}_+^{n_\mathrm{p}}$ denote a constrained minimizer of f,* i.e., *$f(\boldsymbol{x}^{(\star)}) \le f(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in \mathbb{R}_+^{n_\mathrm{p}}$. Then*

$$\langle \nabla f(\boldsymbol{x}^{(\star)}), \, \boldsymbol{D}(\boldsymbol{x} - \boldsymbol{x}^{(\star)}) \rangle \ge 0, \quad \forall \boldsymbol{x} \in \mathbb{R}_+^{n_\mathrm{p}}.$$

Proof:
By the **Karush-Kuhn-Tucker** (**KKT**) conditions,

$$\frac{\partial}{\partial x_j} f(\boldsymbol{x}^{(\star)}) \begin{cases} = 0, & x_j^\star > 0 \\ \ge 0, & x_j^\star = 0, \end{cases} \tag{29.10.1}$$

e,math,min,nonneg,kkt

so $x_j^\star \frac{\partial}{\partial x_j} f(\boldsymbol{x}^{(\star)}) = 0$ and hence $\sum_{j=1}^{n_\mathrm{p}} d_j x_j^\star \frac{\partial}{\partial x_j} f(\boldsymbol{x}^{(\star)}) = 0$. Furthermore, $\boldsymbol{x} \in \mathbb{R}_+^{n_\mathrm{p}}$ implies $x_j \ge 0$, so

$$\sum_{j=1}^{n_\mathrm{p}} d_j x_j \frac{\partial}{\partial x_j} f(\boldsymbol{x}^{(\star)}) = 0,$$

because $\frac{\partial}{\partial x_j} f(\boldsymbol{x}^{(\star)}) \ge 0$. □

s,math,al

29.11 Augmented Lagrangian methods (s,math,al)

This section summarizes the **augmented Lagrangian** method for constrained optimization [20, Ch. 3] [21], considering the case of complex vectors.

Let $\Psi : \mathbb{C}^{n_\mathrm{p}} \mapsto \mathbb{R}$ denote a cost function and consider the constrained minimization problem

e,math,al,p0

$$\arg\min_{\boldsymbol{x} \in \mathbb{C}^{n_\mathrm{p}}} \Psi(\boldsymbol{x}) \text{ sub. to } \boldsymbol{B}\boldsymbol{x} = \boldsymbol{c}, \tag{29.11.1}$$

where $\boldsymbol{B}$ is a $M \times n_\mathrm{p}$ matrix (possibly complex valued) and $\boldsymbol{c} \in \mathbb{C}^M$.

It is easy to verify that (29.11.1) is equivalent to the following **saddle-point problem**:

e,math,al,p1

$$\max_{\boldsymbol{q} \in \mathbb{C}^M} \arg\min_{\boldsymbol{x} \in \mathbb{C}^{n_\mathrm{p}}} \Psi(\boldsymbol{x}) + \mathrm{real}\{\boldsymbol{q}'(\boldsymbol{B}\boldsymbol{x} - \boldsymbol{c})\}, \tag{29.11.2}$$

because if we choose $\boldsymbol{q} = \alpha(\boldsymbol{B}\boldsymbol{x} - \boldsymbol{c})$ for $\alpha \in \mathbb{R}$ then the second term becomes $\alpha \left\|\boldsymbol{B}\boldsymbol{x} - \boldsymbol{c}\right\|^2$ which is unbounded as $\alpha \to \infty$ unless $\boldsymbol{B}\boldsymbol{x} = \boldsymbol{c}$. The vector $\boldsymbol{q}$ is called a **Lagrange multiplier**.

The saddle-point problem (29.11.2) is also equivalent to the following problem:

e,math,al,p2

$$\max_{\boldsymbol{q} \in \mathbb{C}^M} \arg\min_{\boldsymbol{x} \in \mathbb{C}^{n_\mathrm{p}}} L_\mu(\boldsymbol{x}, \boldsymbol{q}), \tag{29.11.3}$$

where the **augmented Lagrangian** is defined by

e,math,al,a1

$$L_\mu(\boldsymbol{x}, \boldsymbol{q}) \triangleq \Psi(\boldsymbol{x}) + \mathrm{real}\{\boldsymbol{q}'(\boldsymbol{B}\boldsymbol{x} - \boldsymbol{c})\} + \frac{\mu}{2} \left\|\boldsymbol{B}\boldsymbol{x} - \boldsymbol{c}\right\|^2, \tag{29.11.4}$$

for any real $\mu > 0$. By **completing the square**, one can rewrite the **augmented Lagrangian** as[2]

e,math,al,a12

$$L_\mu(\boldsymbol{x}, \boldsymbol{q}) = \Psi(\boldsymbol{x}) + \frac{\mu}{2} \left\|\boldsymbol{B}\boldsymbol{x} - \boldsymbol{c} + \frac{1}{\mu}\boldsymbol{q}\right\|^2 - \frac{\mu}{2} \left\|\frac{1}{\mu}\boldsymbol{q}\right\|^2. \tag{29.11.5}$$

Identifying $\boldsymbol{\eta} = \frac{-1}{\mu}\boldsymbol{q}$ we rewrite (29.11.3) as

$$\max_{\boldsymbol{\eta} \in \mathbb{C}^M} \arg\min_{\boldsymbol{x} \in \mathbb{C}^{n_\mathrm{p}}} \tilde{L}_\mu(\boldsymbol{x}, \boldsymbol{\eta}) \text{ where } \tilde{L}_\mu(\boldsymbol{x}, \boldsymbol{\eta}) \triangleq L_\mu(\boldsymbol{x}, -\mu\boldsymbol{\eta}) = \Psi(\boldsymbol{x}) + \frac{\mu}{2} \left\|\boldsymbol{B}\boldsymbol{x} - \boldsymbol{c} - \boldsymbol{\eta}\right\|^2 - \frac{\mu}{2} \left\|\boldsymbol{\eta}\right\|^2.$$

One way to solve this saddle-point problem is to alternate between updating $\boldsymbol{x}$ and updating $\boldsymbol{\eta}$. We update $\boldsymbol{x}$ using

$$\boldsymbol{x}^{(n+1)} = \arg\min_{\boldsymbol{x} \in \mathbb{C}^{n_\mathrm{p}}} \tilde{L}_\mu(\boldsymbol{x}, \boldsymbol{\eta}^{(n)}) = \arg\min_{\boldsymbol{x} \in \mathbb{C}^{n_\mathrm{p}}} \Psi(\boldsymbol{x}) + \frac{\mu}{2} \left\|\boldsymbol{B}\boldsymbol{x} - \boldsymbol{c} - \boldsymbol{\eta}^{(n)}\right\|^2.$$

At first it might seem natural to update $\boldsymbol{\eta}$ using

$$\boldsymbol{\eta}^{(n+1)} = \arg\max_{\boldsymbol{\eta} \in \mathbb{C}^M} \tilde{L}_\mu(\boldsymbol{x}^{(n+1)}, \boldsymbol{\eta}).$$

However, that would lead to $\boldsymbol{\eta}^{(n+1)}$ with infinite norm when $\boldsymbol{B}\boldsymbol{x}^{(n)} \neq \boldsymbol{c}$. Instead we apply a single **gradient ascent** update (using (29.2.16) in the complex case) for the Lagrange multiplier:

$$\boldsymbol{\eta}^{(n+1)} \triangleq \boldsymbol{\eta}^{(n)} + \alpha_n \nabla_{\boldsymbol{\eta}} \tilde{L}_\mu(\boldsymbol{x}^{(n+1)}, \boldsymbol{\eta}^{(n)}) = \boldsymbol{\eta}^{(n)} - \alpha_n \mu \left(\boldsymbol{B}\boldsymbol{x} - \boldsymbol{c}\right).$$

The parameter α_n is a step size that must be chosen appropriately to ensure convergence [20, Ch. 3]. Other methods for updating the Lagrange multiplier $\boldsymbol{\eta}$ have been proposed [23, 24].

s,math,floor

29.12 Special functions (s,math,floor)

The **floor** and **ceiling** operations often are useful for imaging geometry calculations. They are defined as follows:

$$\lfloor x \rfloor \triangleq \max\{n \in \mathbb{Z} : n \leq x\}$$
$$\lceil x \rceil \triangleq \min\{n \in \mathbb{Z} : n \geq x\}.$$

[2] Ramani et al. [22, after (13)] imply that the last term is an irrelevant constant. It is independent of $\boldsymbol{x}$, so irrelevant for updating $\boldsymbol{x}$, but it is essential for the update of $\boldsymbol{\eta}$.

Useful properties of $\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$ include the following:

$$\begin{aligned} x \leq \lceil x \rceil &< 1 + x \\ x - 1 < \lfloor x \rfloor &\leq x \\ \lceil x \rceil \leq \lfloor x + 1 \rfloor &= \lfloor x \rfloor + 1 \end{aligned} \tag{29.12.1}$$

e,math,floor,ceil<=floor+1

$$\begin{aligned} \lfloor -x \rfloor &= -\lceil x \rceil \\ \lceil -x \rceil &= -\lfloor x \rfloor \\ \lceil x + b \rceil - \lfloor x + a \rfloor &\leq 1 + \lceil b - a \rceil . \end{aligned} \tag{29.12.2}$$

e,math,floor,ceil-floor

29.13 Convergence rates of iterations (s,math,rate)

s,math,rate

One can evaluate the **convergence rate** of iterative methods using a variety of metrics.

- When a cost function Ψ has a minimum value $\Psi_* = \min_{\boldsymbol{z}} \Psi(\boldsymbol{x})$, we can examine how quickly $\Psi(\boldsymbol{x}^{(n)}) - \Psi_*$ approaches zero.
- If a cost function Ψ is differentiable, we can examine how quickly a norm of its gradient $\|\nabla \Psi(\boldsymbol{x}^{(n)})\|$ approaches zero.
- If a cost function Ψ has a unique minimizer $\boldsymbol{x}^{(\star)}$, we can examine how quickly $\|\boldsymbol{x}^{(n)} - \boldsymbol{x}^{(\star)}\|$ approaches zero.

The rest of this section focuses on the latter case because it is of the most interest.

We a say a sequence $\{\boldsymbol{x}^{(n)}\}$ **converges** to a limit $\boldsymbol{x}^{(\star)}$ with respect to a **norm** $\|\cdot\|$ if for all $\varepsilon > 0$ there exists a number N_ε such that

$$\|\boldsymbol{x}^{(n)} - \boldsymbol{x}^{(\star)}\| < \varepsilon, \qquad \forall n \geq N_\varepsilon .$$

There are various ways to quantify the **asymptotic convergence rate** of convergent sequences, and these are important for understanding the limiting behavior of optimization methods.

The **quotient convergence factor** of a sequence $\{\boldsymbol{x}^{(n)}\}$ converging to a limit $\boldsymbol{x}^{(\star)}$ is defined by [25, p. 281] [26]

$$Q_1(\{\boldsymbol{x}^{(n)}\}) \triangleq \limsup_{n \to \infty} \frac{\left\|\boldsymbol{x}^{(n+1)} - \boldsymbol{x}^{(\star)}\right\|}{\|\boldsymbol{x}^{(n)} - \boldsymbol{x}^{(\star)}\|} . \tag{29.13.1}$$

e,math,rate,quotient,xn

If this **limit superior** lies in the interval $(0, 1)$ then we say that the sequence $\{\boldsymbol{x}^{(n)}\}$ **converges linearly** to $\boldsymbol{x}^{(\star)}$.

Another measure of convergence rate of a sequence $\{\boldsymbol{x}^{(n)}\}$ converging to $\boldsymbol{x}^{(\star)}$ is the **root convergence factor** defined by [25, p. 288]

$$R_1(\{\boldsymbol{x}^{(n)}\}) \triangleq \limsup_{n \to \infty} \|\boldsymbol{x}^{(n)} - \boldsymbol{x}^{(\star)}\|^{1/n} . \tag{29.13.2}$$

e,math,rate,root,xn

One can show that $0 \leq R_1 \leq 1$. Again, if this limit superior lies in the interval $(0, 1)$ then we say that the sequence **converges linearly** to $\boldsymbol{x}^{(\star)}$.

Unlike the quotient convergence factor Q_1, the root convergence factor R_1 is independent of the norm [25, p. 288]. Furthermore, $R_1 \leq Q_1$ for any norm [25, p. 296].

Often we are more interested in defining the convergence rate of an **iterative process** $\boldsymbol{x}^{(n+1)} = \mathcal{M}(\boldsymbol{x}^{(n)})$, where $\mathcal{M} : \mathbb{R}^{n_{\mathrm{P}}} \to \mathbb{R}^{n_{\mathrm{P}}}$, rather than that of a specific sequence generated by the process [27]. After all, if we happen to initialize the algorithm with $\boldsymbol{x}^{(0)} = \boldsymbol{x}^{(\star)}$ where $\boldsymbol{x}^{(\star)} = \mathcal{M}(\boldsymbol{x}^{(\star)})$ is a **fixed point** of $\mathcal{M}$ then we would converge immediately. So clearly the convergence rate of a particular sequence can depend on the initial condition. To avoid this dependence we examine the **worst case** over all possible initializers (or, more precisely, over all initializers that lead to convergence to $\boldsymbol{x}^{(\star)}$):

$$Q_1(\mathcal{M}) \triangleq \sup_{\boldsymbol{x}^{(0)}} \left\{ Q_1(\{\boldsymbol{x}^{(n)}\}) : \boldsymbol{x}^{(n+1)} = \mathcal{M}(\boldsymbol{x}^{(n)}) \to \boldsymbol{x}^{(\star)} \right\} . \tag{29.13.3}$$

e,math,rate,quotient,M

We call this the **quotient convergence factor** of the iterative process $\mathcal{M}$. Likewise we can define a **root convergence factor** of $\mathcal{M}$ by

$$R_1(\mathcal{M}) \triangleq \sup_{\boldsymbol{x}^{(0)}} \left\{ R_1(\{\boldsymbol{x}^{(n)}\}) : \boldsymbol{x}^{(n+1)} = \mathcal{M}(\boldsymbol{x}^{(n)}) \to \boldsymbol{x}^{(\star)} \right\} . \tag{29.13.4}$$

e,math,rate,root,M

Ostrowski's theorem [25, p. 300] states that if $\mathcal{M}$ is continuous and differentiable on an open set containing $\boldsymbol{x}^{(\star)}$ and

$$\rho(\nabla \mathcal{M}(\boldsymbol{x}^{(\star)})) < 1,$$

where $\nabla \mathcal{M}(\boldsymbol{x})$ is the $n_\mathrm{p} \times n_\mathrm{p}$ matrix with elements

$$[\nabla \mathcal{M}(\boldsymbol{x})]_{k,j} = \frac{\partial}{\partial x_j} [\mathcal{M}(\boldsymbol{x})]_k \,,$$

and where $\rho(\cdot)$ in (27.1.2) denotes the **spectral radius** (largest eigenvalue magnitude) of a square matrix, then the root convergence factor of $\mathcal{M}$ is

e,math,rate,R1,rho

$$R_1(\mathcal{M}) = \rho(\nabla \mathcal{M}(\boldsymbol{x}^{(\star)})). \tag{29.13.5}$$

As a sketch of why this equality holds, for $\boldsymbol{x}^{(n)} \approx \boldsymbol{x}^{(\star)}$:

$$\boldsymbol{x}^{(n+1)} \approx \mathcal{M}(\boldsymbol{x}^{(\star)}) + \nabla \mathcal{M}(\boldsymbol{x}^{(\star)})(\boldsymbol{x}^{(n)} - \boldsymbol{x}^{(\star)}) = \boldsymbol{x}^{(\star)} + \nabla \mathcal{M}(\boldsymbol{x}^{(\star)})(\boldsymbol{x}^{(n)} - \boldsymbol{x}^{(\star)})$$

so

$$\boldsymbol{x}^{(n+1)} - \boldsymbol{x}^{(\star)} \approx \nabla \mathcal{M}(\boldsymbol{x}^{(\star)})(\boldsymbol{x}^{(n)} - \boldsymbol{x}^{(\star)}).$$

Thus the asymptotic rate of convergence of $\boldsymbol{x}^{(n)}$ to $\boldsymbol{x}^{(\star)}$ is governed by the eigenvalues of $\nabla \mathcal{M}(\boldsymbol{x}^{(\star)})$.

x,math,rate,linear

Example 29.13.1 *Consider a sequence generated by the affine recursion*

e,math,rate,M,affine

$$\boldsymbol{x}^{(n+1)} = \mathcal{M}(\boldsymbol{x}^{(n)}) \triangleq \boldsymbol{B}\boldsymbol{x}^{(n)} + \boldsymbol{u} \tag{29.13.6}$$

*where $\boldsymbol{I} - \boldsymbol{B}$ is an invertible matrix and we define $\boldsymbol{x}^{(\star)} = [\boldsymbol{I} - \boldsymbol{B}]^{-1} \boldsymbol{u}$ so that $\boldsymbol{x}^{(n+1)} - \boldsymbol{x}^{(\star)} = \boldsymbol{B}(\boldsymbol{x}^{(n)} - \boldsymbol{x}^{(\star)})$ and hence $\boldsymbol{x}^{(n)} - \boldsymbol{x}^{(\star)} = \boldsymbol{B}^n(\boldsymbol{x}^{(0)} - \boldsymbol{x}^{(\star)})$. If $\rho(\boldsymbol{B}) < 1$ then it is easy to show [25, p. 303] using the **Jordan form** of $\boldsymbol{B}$ that*

$$R_1(\mathcal{M}) = \rho(\boldsymbol{B}).$$

As a more concrete example, the **PGD** iteration $\boldsymbol{x}^{(n+1)} = \boldsymbol{x}^{(n)} - \boldsymbol{P}\nabla\Psi(\boldsymbol{x}^{(n)})$ for the quadratic cost function $\Psi(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}'\boldsymbol{H}\boldsymbol{x} - \boldsymbol{b}'\boldsymbol{x}$ has the form (29.13.6) with $\boldsymbol{B} = \boldsymbol{I} - \boldsymbol{P}\boldsymbol{H}$ and $\boldsymbol{u} = \boldsymbol{P}\boldsymbol{b}$. So the root convergence factor for PGD in this case is $R_1 = \rho(\boldsymbol{I} - \boldsymbol{P}\boldsymbol{H})$.

29.14 Problems (s,math,prob)

s,math,prob

p,math,four,scale

Problem 29.1 *Prove affine scaling property (29.2.5) of the Fourier transform.* *(Need typed.)*

math,lap1,cos,exp,step

Problem 29.2 *Prove the Laplace transform pair for* $\cos(bt)\,\mathrm{e}^{-at}\,\mathrm{step}(t)$ *in §29.4.1.*

math,lap1,sin,exp,step

Problem 29.3 *Prove the Laplace transform pair for* $\sin(bt)\,\mathrm{e}^{-at}\,\mathrm{step}(t)$ *in §29.4.1.*

p,math,lap1,cos,exp

Problem 29.4 *Prove the Laplace transform pair for* $\cos(bt)\,\mathrm{e}^{-a|t|}$ *in §29.4.1*

p,math,lap1,sin,exp

Problem 29.5 *Prove the Laplace transform pair for* $\sin(b|t|)\,\mathrm{e}^{-a|t|}$ *in §29.4.1*

p,math,convex,affine

Problem 29.6 *Prove that if f is convex on $\mathbb{R}^n$ then $g(\boldsymbol{z}) = f(\boldsymbol{M}\boldsymbol{z} + \boldsymbol{b})$ is convex.* *(Need typed.)*

p,math,convex,lp

Problem 29.7 *Prove that $f(x) = |x|^p$ is strictly convex for $p > 1$, for Example 29.9.10.*

math,convex,gp,upper,s

Problem 29.8 *Prove Corollary 29.9.21,* i.e., *(29.9.13).*

p,math,lip,majorize

Problem 29.9 *Prove Corollary 29.9.20,* i.e., *for a twice differentiable, convex cost function Ψ, the Lipschitz condition (29.9.12) is equivalent to $\nabla^2\Psi(\boldsymbol{x}) \preceq \boldsymbol{S}\boldsymbol{S}'$. Hint: use Theorem 27.5.3.*

29.15 Bibliography

luenberger:69

[1] D. G. Luenberger. *Optimization by vector space methods*. New York: Wiley, 1969. URL: http://books.google.com/books?id=lZU0CAH4RccC (cit. on p. 29.9).

boyd:04

[2] S. Boyd and L. Vandenberghe. *Convex optimization*. UK: Cambridge, 2004. URL: http://web.stanford.edu/~boyd/cvxbook.html (cit. on pp. 29.10, 29.11, 29.12).

wiesel:12:gca
[3] A. Wiesel. "Geodesic convexity and covariance estimation." In: *IEEE Trans. Sig. Proc.* 60.12 (Dec. 2012), 6182–9. DOI: 10.1109/TSP.2012.2218241 (cit. on p. 29.10).

bertsekas:82
[4] D. P. Bertsekas. *Constrained optimization and Lagrange multiplier methods*. New York: Academic-Press, 1982. URL: http://www.athenasc.com/ (cit. on p. 29.10).

horn:85
[5] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge: Cambridge Univ. Press, 1985 (cit. on p. 29.11).

polyak:87
[6] B. T. Polyak. *Introduction to optimization*. New York: Optimization Software Inc, 1987 (cit. on pp. 29.12, 29.13, 29.15).

nesterov:04
[7] Y. Nesterov. *Introductory lectures on convex optimization: A basic course*. Kluwer, 2004. DOI: 10.1007/978-1-4419-8853-9 (cit. on pp. 29.12, 29.13).

giselsson:14:mar
[8] P. Giselsson and S. Boyd. "Monotonicity and restart in fast gradient methods." In: *Proc. Conf. Decision and Control*. 2014, 5058–63. DOI: 10.1109/CDC.2014.7040179 (cit. on p. 29.13).

moreau:62:fcd
[9] J-J. Moreau. "Fonctions convexes duales et points proximaux dans un espace hilbertien." In: *C. R. Acad. Sci. Paris* 255 (1962), 2897–289 (cit. on p. 29.14).

moreau:65:ped
[10] J. J. Moreau. "Proximité et dualité dans un espace hilbertien." In: *Bulletin de la Société Mathématique de France* 93 (1965), 273–99. URL: http://www.numdam.org/item?id=BSMF_1965__93__273_0 (cit. on p. 29.14).

combettes:05:srb
[11] P. Combettes and V. Wajs. "Signal recovery by proximal forward-backward splitting." In: *siam-jmms* 4.4 (2005), 1168–200. DOI: 10.1137/050626090 (cit. on p. 29.14).

parikh:13:pa
[12] N. Parikh and S. Boyd. "Proximal algorithms." In: *Found. Trends in Optimization* 1.3 (2013), 123–231. DOI: 10.1561/2400000003 (cit. on p. 29.14).

beck:09:afi
[13] A. Beck and M. Teboulle. "A fast iterative shrinkage-thresholding algorithm for linear inverse problems." In: *SIAM J. Imaging Sci.* 2.1 (2009), 183–202. DOI: 10.1137/080716542 (cit. on p. 29.14).

beck:09:fgb
[14] A. Beck and M. Teboulle. "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems." In: *IEEE Trans. Im. Proc.* 18.11 (Nov. 2009), 2419–34. DOI: 10.1109/TIP.2009.2028250 (cit. on p. 29.14).

beck:12:saf
[15] A. Beck and M. Teboulle. "Smoothing and first order methods: A unified framework." In: *SIAM J. Optim.* 22.2 (2012), 557–80. DOI: 10.1137/100818327 (cit. on p. 29.14).

bauschke:11:caa
[16] H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. Springer, 2011. DOI: 10.1007/978-1-4419-9467-7 (cit. on p. 29.14).

sion:58:ogm
[17] M. Sion. "On general minimax theorems." In: *Pacific J. Math.* 8.1 (1958), 171–6. DOI: 10.2140/pjm.1958.8.171 (cit. on p. 29.15).

komiya:88:epf
[18] H. Komiya. "Elementary proof for Sion's minimax theorem." In: *Kodai Mathematical J.* 11.1 (1988), 5–7. DOI: 10.2996/kmj/1138038812 (cit. on p. 29.15).

mourya:15:alw
[19] R. Mourya et al. "Augmented Lagrangian without alternating directions: practical algorithms for inverse problems in imaging." In: *Proc. IEEE Intl. Conf. on Image Processing*. 2015, 1205–9. DOI: 10.1109/ICIP.2015.7350991.

glowinski:89
[20] R. Glowinski and P. L. Tallec. *Augmented Lagrangian and operator-splitting methods in nonlinear mechanics*. Soc. Indust. Appl. Math., 1989. DOI: 10.1137/1.9781611970838.ch3 (cit. on p. 29.16).

birgin:14
[21] E. G. Birgin and J. M. Martinez. *Practical augmented Lagrangian methods for constrained optimization*. ISBN 978-1-611973-35-8. Soc. Indust. Appl. Math., 2014. URL: http://www.ec-securehost.com/SIAM/FA10.html (cit. on p. 29.16).

ramani:11:pmi
[22] S. Ramani and J. A. Fessler. "Parallel MR image reconstruction using augmented Lagrangian methods." In: *IEEE Trans. Med. Imag.* 30.3 (Mar. 2011), 694–706. DOI: 10.1109/TMI.2010.2093536 (cit. on p. 29.16).

eckstein:94:pad
[23] J. Eckstein. "Parallel alternating direction multiplier decomposition of convex programs." In: *J. Optim. Theory Appl.* 80.1 (Jan. 1994), 39–62. DOI: 10.1007/BF02196592 (cit. on p. 29.16).

boyd:10:doa
[24] S. Boyd et al. "Distributed optimization and statistical learning via the alternating direction method of multipliers." In: *Found. & Trends in Machine Learning* 3.1 (2010), 1–122. DOI: 10.1561/2200000016 (cit. on p. 29.16).

ortega:70
[25] J. M. Ortega and W. C. Rheinboldt. *Iterative solution of nonlinear equations in several variables*. New York: Academic, 1970. DOI: 10.1137/1.9780898719468 (cit. on pp. 29.17, 29.18).

potra:89:oqo
[26] F. A. Potra. "On Q-order and R-order of convergence." In: *J. Optim. Theory Appl.* 63.3 (1989), 415–31. DOI: 10.1007/BF00939805 (cit. on p. 29.17).

dubeau:14:fpa
[27] F. Dubeau and C. Gnang. "Fixed point and Newton's methods for solving a nonlinear equation: from linear to high-order convergence." In: *SIAM Review* 56.4 (Dec. 2014), 691–708. DOI: 10.1137/130934799 (cit. on p. 29.17).