abaisero.sty

Andrea Baisero

August 13, 2024

1 Options

Options are processes left-to-right. If no options are provided, or none are enabled by the end of the processing, then by default they are all considered to be enabled.

Option	Description
`all`	Enable all commands
`(no-)math`	Disable/enable mathematical commands
`(no-)linalg`	Disable/enable linear algebra commands
`(no-)optim`	Disable/enable optimization commands
`(no-)stats`	Disable/enable statistics commands
`(no-)dists`	Disable/enable distributions commands
`(no-)ml`	Disable/enable machine learning commands
`(no-)rl`	Disable/enable reinforcement learning commands
`(no-)marl`	Disable/enable multi-agent reinforcement learning commands
`(no-)theorem`	Disable/enable theorem commands
`(no-)misc`	Disable/enable miscellanea commands

2 Commands

Option [math]

Symbol	Command	Description	Example
$\mathbb{N}$	`\naturalset`	the set of natural numbers	$\mathbb{N} \doteq \{1, 2, 3, \dots\}$
$\mathbb{Z}$	`\integerset`	the set of integer numbers	$\mathbb{Z} \doteq \{0, 1, -1, 2, -3, \dots\}$
$\mathbb{R}$	`\realset`	the set of real numbers	$\sqrt{2} \in \mathbb{R}$
*	`\kstar`	the Kleene star operator	$\mathcal{X}^* \doteq \bigcup_{k=0}^{\infty} \mathcal{X}^k$
$^+$	`\kplus`	the Kleene plus operator	$\mathcal{X}^+ \doteq \bigcup_{k=1}^{\infty} \mathcal{X}^k$
softmax	`\softmax`	a.k.a. logsumexp, realsoftmax[1]	$\operatorname{softmax}(x_1, \dots, x_n) \doteq \log \sum_i \exp(x_i)$
softmin	`\softmin`		$\operatorname{softmin}(x_1, \dots, x_n) \doteq -\log \sum_i \exp(-x_i)$
softargmax	`\softargmax`	a.k.a. softmax[1]	$\operatorname{softargmax}(x_1, \dots, x_n)_i \doteq \frac{\exp(x_i)}{\sum_k \exp(x_k)}$
sign	`\sign`		$x = \operatorname{sign} x \cdot \lvert x \rvert$

[1] The functions that in this document are called "softmax" and "softargmax" are poorly and inaccurately named in the broader math and ML fields (see `https://en.wikipedia.org/wiki/Softmax_function` and `https://en.wikipedia.org/wiki/LogSumExp`). Rather than stick to the more common naming conventions, I opt to rename the functions more accurately to appropriately reflect their actual properties. In any document where I would use these functions, I would need to define them anyway, so the risk of misunderstandings are minimal.

Option [linalg]

Symbol	Command	Description	Example
diag	\diag		
rank	\rank		
tr	\trace		$\mathrm{tr}(M) \doteq \sum_{i=1}^{n} M_{ii}$
col	\colspace		
ker	\nullspace	Nullspace (a.k.a kernel) of a linear mapping	
span	\spanspace		
$^\top$	\T	Transpose superscript	symmetric $M \implies M = M^\top$
$^{-1}$	\I	Inverse superscript	invertible $M \implies MM^{-1} = I$
$^+$	\PI	Pseudo-inverse superscript	$MM^+M = M$
$^{-\top}$	\IT	Inverse transpose superscript	$M^{-\top} = (M^{-1})^\top = (M^\top)^{-1}$
$^{+\top}$	\PIT	Pseudo-inverse transpose superscript	$M^{+\top} = (M^+)^\top = (M^\top)^+$

Option [optim]

Symbol	Command	Description	Example
argmax	\argmax		$\mathrm{argmax}_a \, Q^\pi(s,a)$
argmin	\argmin		$\theta^* \doteq \mathrm{argmin}_\theta \, \mathcal{L}(\theta)$
*	\opt	Optimality superscript	$\pi^*(s) = \mathrm{argmax}_a \, Q^*(s,a)$

Option [stats]

Symbol	Command	Description	Example
$\perp$	\indep	Independence	$X \perp Y \mid Z$
$\mathbb{C}$	\Cov	Covariance	$\mathbb{C}(x,y) = \mathbb{E}[xy] - \mathbb{E}[x]\,\mathbb{E}[y]$
$\mathbb{H}$	\Ent	Entropy	$\mathbb{H}[x] = -\,\mathbb{E}[\log \Pr(x)]$
$\mathbb{E}$	\Exp	Expectation	$\mathbb{E}[f(x)] = \sum_x \Pr(x) f(x)$
$\mathbb{I}$	\Ind	Indicator function	$\Pr(x=0) = \mathbb{E}[\mathbb{I}[x=0]]$
KL	\KL	KL-divergence	$\mathrm{KL}(p \,||\, q) \doteq \mathbb{E}_{x\sim p}[\log p(x) - \log q(x)]$
D_{KL}	\DKL	KL-divergence (alternative)	
$\mathbb{I}$	\MI	Mutual Information	
$\mathbb{B}$	\Bias	Bias	$\mathbb{B}\left[\hat{f}(x)\right]$ is the bias of estimator $\hat{f}$
$\mathbb{V}$	\Var	Variance	$\mathbb{V}\left[\hat{f}(x)\right] = \mathbb{E}\left[\hat{f}(x)^2\right] - \mathbb{E}\left[\hat{f}(x)\right]^2$

Option [dists]

Symbol	Command	Description
Categorical	\Categorical	Categorical
Dirichlet	\Dirichlet	Dirichlet
Geometric	\Geometric	Geometric
Normal	\Normal	Normal
Uniform	\Uniform	Uniform

Option [ml]

Symbol	Command	Description	Example
$\mathcal{D}$	\data	Data set	$\mathcal{D} \doteq \{(x_i, y_i)\}_{i=1}^N$
$\mathcal{L}$	\loss	Loss function	$\mathcal{L}(\theta; x, y) = \frac{1}{2}\|y - f(x;\theta)\|^2$
nll	\nll	Neg-log-likelihood	$\text{nll}(x) \doteq -\log \Pr(x)$
MSE	\mse	Mean-squared-error	

Option [rl]

Symbol	Command	Description
$\mathcal{A}$	\aset	Action set
$\mathcal{B}$	\bset	Belief set
$\mathcal{H}$	\hset	History set
$\mathcal{O}$	\oset	Observation set
$\mathcal{R}$	\rset	Reward set
$\mathcal{S}$	\sset	State set
D	\dfn	Dynamics function
G	\gfn	Generative function
O	\ofn	Observation function
R	\rfn	Reward function
T	\tfn	Transition function
ε	\nohistory	Empty history
π	\policy	policy
Q^π	\qpolicy	Q policy values
$\hat{Q}$	\qmodel	Parametric model
V^π	\vpolicy	V policy values
$\hat{V}$	\vmodel	Parametric model
A^π	\apolicy	A policy values
$\hat{A}$	\amodel	Parametric model
U^π	\upolicy	U policy values
$\hat{U}$	\umodel	Parametric model

Option [marl]

Symbol	Command	Description
$\bar{\mathcal{H}}$	\hsset	Joint history set
$\bar{\mathcal{A}}$	\asset	Joint action set
$\bar{\mathcal{O}}$	\osset	Joint observation set
$\bar{h}$	\hs	Joint history
$\bar{a}$	\as	Joint action
$\bar{o}$	\os	Joint observation
$\bar{\pi}$	\policies	Joint policy
$Q^{\bar{\pi}}$	\qpolicies	Q joint-policy values
$V^{\bar{\pi}}$	\vpolicies	V joint-policy values
$A^{\bar{\pi}}$	\apolicies	A joint-policy values
$U^{\bar{\pi}}$	\upolicies	U joint-policy values

Option [theorem]

Symbol	Command	Description
	\begin{definition}	
	\begin{example}	
	\begin{axiom}	
	\begin{conjecture}	
	\begin{proposition}	
	\begin{lemma}	
	\begin{theorem}	
	\begin{corollary}	
	\begin{generalization}	

Option [misc]

Symbol	Command	Description
$\dagger$	\D	Dagger superscript
(k)	\iter{k}	Superscript indicating iteration