`latex-math` Macros

2021-06-07

Latex macros like `\frac{#1}{#2}` with arguments are displayed as $\frac{\#1}{\#2}$.

basic-math.tex

Macro	Notation	Comment
`\N`	$\mathbb{N}$	N defined by "siunitx" (which we use), for "NEWTON"
`\N`	$\mathbb{N}$	
`\Z`	$\mathbb{Z}$	Z, integers
`\Q`	$\mathbb{Q}$	Q, rationals
`\R`	$\mathbb{R}$	R, reals
`\C`	$\mathbb{C}$	C, complex
`\C`	$\mathbb{C}$	
`\continuous`	$\mathcal{C}$	C, space of continuous functions
`\M`	$\mathcal{M}$	machine numbers
`\epsm`	ϵ_m	maximum error
`\xt`	$\tilde{x}$	x tilde
`\sign`	sign	sign, signum
`\I`	$\mathbb{I}$	I, indicator
`\Ind`	$\mathbb{1}$	1, indicator
`\order`	$\mathcal{O}$	O, order
`\fp`	$\frac{\partial \#1}{\partial \#2}$	partial derivative
`\pd`	$\frac{\partial \#1}{\partial \#2}$	partial derivative
`\sumin`	$\sum\limits_{i=1}^n$	summation from i=1 to n
`\sumjp`	$\sum\limits_{j=1}^p$	summation from j=1 to p
`\sumik`	$\sum\limits_{i=1}^k$	summation from i=1 to k
`\sumkg`	$\sum\limits_{k=1}^g$	summation from k=1 to g
`\sumjg`	$\sum\limits_{j=1}^g$	summation from j=1 to g
`\meanin`	$\frac{1}{n}\sum\limits_{i=1}^n$	mean from i=1 to n
`\meankg`	$\frac{1}{g}\sum\limits_{k=1}^g$	mean from k=1 to g
`\prodin`	$\prod\limits_{i=1}^n$	product from i=1 to n
`\prodkg`	$\prod\limits_{k=1}^g$	product from k=1 to g
`\prodjp`	$\prod\limits_{j=1}^p$	product from j=1 to p
`\one`	$\mathbf{1}$	1, unitvector

<table>
<tr><td>\zero</td><td>$\mathbf{0}$</td><td>0-vector</td></tr>
<tr><td>\id</td><td>$\boldsymbol{I}$</td><td>I, identity</td></tr>
<tr><td>\diag</td><td>diag</td><td>diag, diagonal</td></tr>
<tr><td>\trace</td><td>tr</td><td>tr, trace</td></tr>
<tr><td>\spn</td><td>span</td><td>span</td></tr>
<tr><td>\scp</td><td>$\langle \#1, \#2 \rangle$</td><td><.,.>, scalarproduct</td></tr>
<tr><td>\Amat</td><td>$\mathbf{A}$</td><td>matrix A</td></tr>
<tr><td>\xv</td><td>$\mathbf{x}$</td><td>vector x (bold)</td></tr>
<tr><td>\xtil</td><td>$\tilde{\mathbf{x}}$</td><td>vector x-tilde (bold)</td></tr>
<tr><td>\xb</td><td>$\mathbf{x}$</td><td>WE SHOULD NOT USE THIS</td></tr>
<tr><td>\yv</td><td>$\mathbf{y}$</td><td>vector y (bold)</td></tr>
<tr><td>\Deltab</td><td>$\boldsymbol{\Delta}$</td><td>error term for vectors</td></tr>
<tr><td>\P</td><td>$\mathbb{P}$</td><td>P, probability</td></tr>
<tr><td>\E</td><td>$\mathbb{E}$</td><td>E, expectation</td></tr>
<tr><td>\var</td><td>Var</td><td>Var, variance</td></tr>
<tr><td>\cov</td><td>Cov</td><td>Cov, covariance</td></tr>
<tr><td>\corr</td><td>Corr</td><td>Corr, correlation</td></tr>
<tr><td>\normal</td><td>$\mathcal{N}$</td><td>N of the normal distribution</td></tr>
<tr><td>\iid</td><td>$\stackrel{i.i.d}{\sim}$</td><td>dist with i.i.d superscript</td></tr>
<tr><td>\distas</td><td>$\stackrel{\#1}{\sim}$</td><td>... is distributed as ...</td></tr>
<tr><td>\ind</td><td>$\perp$</td><td>__|__, ... is independent of ...</td></tr>
</table>

basic-ml.tex

Macro	Notation	Comment
`\Xspace`	$\mathcal{X}$	X, input space
`\Yspace`	$\mathcal{Y}$	Y, output space
`\allDatasets`	$\mathbb{D}$	The set of all datasets
`\allDatasetsn`	$\mathbb{D}_n$	The set of all datasets of size n
`\defAllDatasetsn`	$(\mathcal{X} \times \mathcal{Y})^n$	Def. of the set of all datasets of size n
`\defAllDatasets`	$\bigcup_{n \in \mathbb{N}} (\mathcal{X} \times \mathcal{Y})^n$	Def. of the set of all datasets
`\nset`	$\{1, \ldots, n\}$	set from 1 to n
`\pset`	$\{1, \ldots, p\}$	set from 1 to p
`\gset`	$\{1, \ldots, g\}$	set from 1 to g
`\Pxy`	$\mathbb{P}_{xy}$	P_xy
`\Exy`	$\mathbb{E}_{xy}$	E_xy: Expectation over random variables xy
`\xy`	$(\mathbf{x}, y)$	observation (x, y)
`\xvec`	$(x_1, \ldots, x_p)^T$	(x1, ..., xp)
`\Xmat`	$\mathbf{X}$	Design matrix
`\D`	$\mathcal{D}$	D, data
`\ydat`	$\mathbf{y}$	y (bold), vector of outcomes
`\yvec`	$\left(y^{(1)}, \ldots, y^{(n)}\right)^T$	(y1, ..., yn), vector of outcomes
`\xi`	$\mathbf{x}^{(\#1)}$	x^i, i-th observed value of x
`\yi`	$y^{(\#1)}$	y^i, i-th observed value of y
`\xyi`	$\left(\mathbf{x}^{(\#1)}, y^{(\#1)}\right)$	(x^i, y^i), i-th observation
`\xivec`	$\left(x_1^{(i)}, \ldots, x_p^{(i)}\right)^T$	(x1^i, ..., xp^i), i-th observation vector
`\xj`	$\mathbf{x}_j$	x_j, j-th feature
`\xij`	$x_j^{(i)}$	x^i_j, j-th feature value of i-th observation
`\xjvec`	$\left(x_j^{(1)}, \ldots, x_j^{(n)}\right)^T$	(x^1_j, ..., x^n_j), j-th feature vector
`\Dtrain`	$\mathcal{D}_{\text{train}}$	D_train, training set
`\Dtest`	$\mathcal{D}_{\text{test}}$	D_test, test set
`\phiv`	ϕ	Basis transformation function phi
`\phixi`	$\phi^{(i)}$	Basis transformation of xi: phi^i := phi(xi)
`\preimageInducer`	$\left(\bigcup_{n \in \mathbb{N}} (\mathcal{X} \times \mathcal{Y})^n\right) \times \mathbf{\Lambda}$	Set of all datasets times the hyperparameter space
`\preimageInducerShort`	$\mathbb{D} \times \mathbf{\Lambda}$	Set of all datasets times the hyperparameter space
`\inducer`	$\mathcal{I}$	Inducer, inducing algorithm, learning algorithm
`\ftrue`	f_{true}	True underlying function (if a statistical model is assumed)
`\ftruex`	$f_{\text{true}}(\mathbf{x})$	True underlying function (if a statistical model is assumed)
`\fx`	$f(\mathbf{x})$	f(x), continuous prediction function
`\Hspace`	$\mathcal{H}$	hypothesis space where f is from
`\fix`	$f_i(\mathbf{x})$	f_i(x), discriminant component function
`\fjx`	$f_j(\mathbf{x})$	f_j(x), discriminant component function
`\fkx`	$f_k(\mathbf{x})$	f_k(x), discriminant component function
`\fgx`	$f_g(\mathbf{x})$	f_g(x), discriminant component function
`\fh`	$\hat{f}$	f hat, estimated prediction function
`\fxh`	$\hat{f}(\mathbf{x})$	fhat(x)
`\fxt`	$f(\mathbf{x} \mid \boldsymbol{\theta})$	f(x | theta)
`\fxi`	$f\left(\mathbf{x}^{(i)}\right)$	f(x^(i))
`\fxih`	$\hat{f}\left(\mathbf{x}^{(i)}\right)$	f(x^(i))
`\fxit`	$f\left(\mathbf{x}^{(i)} \mid \boldsymbol{\theta}\right)$	f(x^(i) | theta)
`\fhD`	$\hat{f}_{\mathcal{D}}$	fhat_D, estimate of f based on D
`\fhDtrain`	$\hat{f}_{\mathcal{D}_{\text{train}}}$	fhat_Dtrain, estimate of f based on D
`\fbayes`	f^*	The Bayes optimal model

`\fxbayes`	$f^*(\mathbf{x})$	The Bayes optimal model
`\hx`	$h(\mathbf{x})$	h(x), discrete prediction function
`\hxv`	$h(\mathbf{x})$	h(x), discrete prediction function with x (vector) as input
`\hh`	$\hat{h}$	h hat
`\hxh`	$\hat{h}(\mathbf{x})$	hhat(x)
`\hxt`	$h(\mathbf{x}\mid\boldsymbol{\theta})$	h(x | theta)
`\hxi`	$h\left(\mathbf{x}^{(i)}\right)$	h(x^(i))
`\hxit`	$h\left(\mathbf{x}^{(i)}\mid\boldsymbol{\theta}\right)$	h(x^(i) | theta)
`\hbayes`	h^*	The Bayes optimal model
`\hxbayes`	$h^*(\mathbf{x})$	The Bayes optimal model
`\yh`	$\hat{y}$	yhat for prediction of target
`\yih`	$\hat{y}^{(i)}$	yhat^(i) for prediction of ith targiet
`\thetah`	$\hat{\boldsymbol{\theta}}$	
`\thetab`	$\boldsymbol{\theta}$	theta vector
`\thetabh`	$\hat{\boldsymbol{\theta}}$	theta vector
`\thetat`	$\boldsymbol{\theta}^{[t]}$	theta^[t] in optimization
`\thetatn`	$\boldsymbol{\theta}^{[t+1]}$	theta^[t+1] in optimization
`\pdf`	p	p
`\pdfx`	$p(\mathbf{x})$	p(x)
`\pixt`	$\pi(\mathbf{x}\mid\boldsymbol{\theta})$	pi(x|theta), pdf of x given theta
`\pixit`	$\pi\left(\mathbf{x}^{(i)}\mid\boldsymbol{\theta}\right)$	pi(x^i|theta), pdf of x given theta
`\pixii`	$\pi\left(\mathbf{x}^{(i)}\right)$	pi(x^i), pdf of i-th x
`\pdfxy`	$p(\mathbf{x},y)$	p(x, y)
`\pdfxyt`	$p(\mathbf{x},y\mid\boldsymbol{\theta})$	p(x, y | theta)
`\pdfxyit`	$p\left(\mathbf{x}^{(i)},y^{(i)}\mid\boldsymbol{\theta}\right)$	p(x^(i), y^(i) | theta)
`\pdfxyk`	$p(\mathbf{x}\mid y=k)$	p(x | y = k)
`\pdfxyj`	$p(\mathbf{x}\mid y=j)$	p(x | y = j)
`\lpdfxyk`	$\log p(\mathbf{x}\mid y=k)$	log p(x | y = k)
`\pdfxiyk`	$p\left(\mathbf{x}^{(i)}\mid y=k\right)$	p(x^i | y = k)
`\pik`	π_k	pi_k, prior
`\lpik`	$\log\pi_k$	log pi_k, log of the prior
`\pit`	$\pi(\boldsymbol{\theta})$	Prior probability of parameter theta
`\post`	$\mathbb{P}(y=1\mid\mathbf{x})$	P(y = 1 | x), post. prob for y=1
`\pix`	$\pi(\mathbf{x})$	
`\pibayes`	π^*	The Bayes optimal model
`\pixbayes`	$\pi^*(\mathbf{x})$	The Bayes optimal model
`\postk`	$\mathbb{P}(y=k\mid\mathbf{x})$	P(y = k | y), post. prob for y=k
`\pikx`	$\pi_k(\mathbf{x})$	pi_k(x), P(y = k | x)
`\pikxt`	$\pi_k(\mathbf{x}\mid\boldsymbol{\theta})$	pi_k(x | theta), P(y = k | x, theta)
`\pijx`	$\pi_j(\mathbf{x})$	pi_j(x), P(y = j | x)
`\pigx`	$\pi_g(\mathbf{x})$	pi_g(x), P(y = g | x)
`\pdfygxt`	$p(y\mid\mathbf{x},\boldsymbol{\theta})$	p(y | x, theta)
`\pdfyigxit`	$p\left(y^{(i)}\mid\mathbf{x}^{(i)},\boldsymbol{\theta}\right)$	p(y^i |x^i, theta)
`\lpdfygxt`	$\log p(y\mid\mathbf{x},\boldsymbol{\theta})$	log p(y | x, theta)
`\lpdfyigxit`	$\log p\left(y^{(i)}\mid\mathbf{x}^{(i)},\boldsymbol{\theta}\right)$	log p(y^i |x^i, theta)
`\pih`	$\hat{\pi}$	pi hat, estimated
`\pixh`	$\hat{\pi}(\mathbf{x})$	pi(x) hat, P(y = 1 | x) hat
`\pikxh`	$\hat{\pi}_k(\mathbf{x})$	pi_k(x) hat, P(y = k | x) hat
`\pixih`	$\hat{\pi}(\mathbf{x}^{(i)})$	pi(x^(i)) with hat
`\pikxih`	$\hat{\pi}_k(\mathbf{x}^{(i)})$	pi_k(x^(i)) with hat
`\eps`	ϵ	residual, stochastic
`\epsi`	$\epsilon^{(i)}$	epsilon^i, residual, stochastic
`\res`	r	residual, empirical

\resi	$r^{(i)}$	epsilon^i, residual, empirical
\epsh	$\hat{\epsilon}$	residual, estimated
\yf	$yf(\mathbf{x})$	y f(x), margin
\yfi	$y^{(i)}f\left(\mathbf{x}^{(i)}\right)$	y^i f(x^i), margin
\Sigmah	$\hat{\Sigma}$	estimated covariance matrix
\Sigmahj	$\hat{\Sigma}_j$	estimated covariance matrix for the j-th class
\Lyf	$L\left(y, f\right)$	L(y, f), loss function
\Lxy	$L\left(y, f(\mathbf{x})\right)$	L(y, f(x)), loss function
\Lxyi	$L\left(y^{(i)}, f\left(\mathbf{x}^{(i)}\right)\right)$	L(y^i, f(x^i))
\Lxyt	$L\left(y, f(\mathbf{x} \mid \boldsymbol{\theta})\right)$	L(y, f(x | theta))
\Lxyit	$L\left(y^{(i)}, f\left(\mathbf{x}^{(i)} \mid \boldsymbol{\theta}\right)\right)$	L(y^i, f(x^i | theta)
\Lxyitt	$L\left(y^{(i)}, f\left(\mathbf{x} \mid \boldsymbol{\theta}^{[t]}\right)\right)$	L(y^i, f(x^i | theta^[t])
\Lxym	$L\left(y^{(i)}, f\left(\tilde{\boldsymbol{x}}^{(i)} \mid \boldsymbol{\theta}\right)\right)$	L(y^i, f(tilde(x)^i | theta),
\Lpixy	$L\left(y, \pi(\mathbf{x})\right)$	L(y, pi(x)), loss function
\Lpixyi	$L\left(y^{(i)}, \pi\left(\mathbf{x}^{(i)}\right)\right)$	L(y^i, pi(x^i))
\Lpixyt	$L\left(y, \pi(\mathbf{x} \mid \boldsymbol{\theta})\right)$	L(y, pi(x | theta))
\Lpixyit	$L\left(y^{(i)}, \pi\left(\mathbf{x}^{(i)} \mid \boldsymbol{\theta}\right)\right)$	L(y^i, pi(x^i | theta)
\Lhxy	$L\left(y, h(\mathbf{x})\right)$	L(y, h(x)), loss function on discrete classes
\Lr	$L\left(r\right)$	L(r), loss function defined on the residual (regression) / margin (class
\risk	$\mathcal{R}$	R, risk
\riskf	$\mathcal{R}(f)$	R(f), risk
\riskt	$\mathcal{R}(\boldsymbol{\theta})$	R(theta), risk
\riske	$\mathcal{R}_{\text{emp}}$	R_emp, empirical risk (without factor 1 / n
\riskeb	$\bar{\mathcal{R}}_{\text{emp}}$	R_emp, empirical risk with factor 1 / n
\riskef	$\mathcal{R}_{\text{emp}}(f)$	R_emp(f)
\risket	$\mathcal{R}_{\text{emp}}(\boldsymbol{\theta})$	R_emp(theta)
\riskr	$\mathcal{R}_{\text{reg}}$	R_reg, regularized risk
\riskrt	$\mathcal{R}_{\text{reg}}(\boldsymbol{\theta})$	R_reg(theta)
\riskrf	$\mathcal{R}_{\text{reg}}(f)$	R_reg(f)
\riskrth	$\hat{\mathcal{R}}_{\text{reg}}(\boldsymbol{\theta})$	hat R_reg(theta)
\risketh	$\hat{\mathcal{R}}_{\text{emp}}(\boldsymbol{\theta})$	hat R_emp(theta)
\riskbayes	$\mathcal{R}^*$	
\LL	$\mathcal{L}$	L, likelihood
\LLt	$\mathcal{L}(\boldsymbol{\theta})$	L(theta), likelihood
\ll	ℓ	l, log-likelihood
\llt	$\ell(\boldsymbol{\theta})$	l(theta), log-likelihood
\LS	$\mathfrak{L}$	????????????
\TS	$\mathfrak{T}$	??????????????
\errtrain	$\text{err}_{\text{train}}$	training error
\errtest	err_{test}	training error
\errexp	$\overline{\text{err}_{\text{test}}}$	training error
\GEf	$GE\left(\hat{f}\right)$	Generalization error of a fitted model
\GEind	$GE_n\left(\mathcal{I}_{L,O}\right)$	Generalization error of a fitted model
\GE	$GE_n\left(\hat{f}_{\#1}\right)$	Generalization error GE
\GEh	$\widehat{GE}_{\#1}$	Estimated train error
\GED	$GE_n\left(\hat{f}_{\mathcal{D}}\right)$	Generalization error GE
\EGEn	EGE_n	Generalization error GE
\EDn	$\mathbb{E}_{\|D\|=n}$	Generalization error GE
\costs	$\mathcal{C}$	costs
\Celite	θ^*	elite configurations
\instances	$\mathcal{I}$	sequence of instances
\budget	$\mathcal{B}$	computational budget

`\np`	n_+	no. of positive instances
`\nn`	n_-	no. of negative instances
`\rn`	π_-	proportion negative instances
`\rp`	π_+	proportion negative instances
`\tp`	$\#\mathrm{TP}$	
`\fap`	$\#\mathrm{FP}$	fp taken for partial derivs
`\tn`	$\#\mathrm{TN}$	
`\fan`	$\#\mathrm{FN}$	

ml-automl.tex

Macro	Notation	Comment
\lambdav	$\boldsymbol{\lambda}$	lambda vector

ml-bagging.tex

Macro	Notation	Comment
`\bl`	$b^{[\#1]}(\mathbf{x})$	baselearner with argument for m
`\blm`	$b^{[m]}(\mathbf{x})$	baselearner without argument for m
`\blmh`	$\hat{b}^{[m]}(\mathbf{x})$	estimated base learner
`\fM`	$f^{[M]}(\mathbf{x})$	ensembled predictor
`\fMh`	$\hat{f}^{[M]}(\mathbf{x})$	estimated ensembled predictor
`\ambifM`	$\Delta\left(f^{[M]}(\mathbf{x})\right)$	ambiguity/instability of ensemble

ml-bayesopt.tex

Macro	Notation	Comment
`\minit`	m_{init}	Size of the initial design
`\lambdai`	$\lambda^{[i]}$	input for black box optimization
`\lambdaopt`	λ^*	Minimum of the black box function Psi
`\metadata`	$\left\{\left(\lambda^{[i]}, \Psi^{[i]}\right)\right\}$	Metadata for the Gaussian process
`\lambdavec`	$\left(\lambda^{[1]}, \dots, \lambda^{[m_{\text{init}}]}\right)$	Vector of different inputs
`\lambdab`	$\boldsymbol{\lambda}$	input configuration
`\gp`	$\mathcal{GP}\left(m(x), k\left(x, x'\right)\right)$	Gaussian Process

ml-boosting.tex

Macro	Notation	Comment
\fm	$f^{[m]}$	prediction in iteration m
\fmh	$\hat{f}^{[m]}$	prediction in iteration m
\fmd	$f^{[m-1]}$	prediction m-1
\fmdh	$\hat{f}^{[m-1]}$	prediction m-1
\bmm	$b^{[m]}$	basemodel m
\bmmxth	$b(\mathbf{x}, \boldsymbol{\theta}^{[m]})$	basemodel of x and theta m
\bmmh	$\hat{b}^{[m]}$	basemodel m with hat
\betam	$\beta^{[m]}$	weight of basemodel m
\betamh	$\hat{\beta}^{[m]}$	weight of basemodel m with hat
\betai	$\beta^{[\#1]}$	weight of basemodel with argument for m
\errm	$\text{err}^{[m]}$	weighted in-sample misclassification rate
\wm	$w^{[m]}$	weight vector of basemodel m
\wmi	$w^{[m](i)}$	weight of obs i of basemodel m
\thetam	$\boldsymbol{\theta}^{[m]}$	parameters of basemodel m
\thetamh	$\hat{\boldsymbol{\theta}}^{[m]}$	parameters of basemodel m with hat
\rmm	$\tilde{r}^{[m]}$	pseudo residuals
\rmi	$\tilde{r}^{[m](i)}$	pseudo residuals
\Rtm	$R_t^{[m]}$	terminal-region
\Tm	$T^{[m]}$	
\ctm	$c_t^{[m]}$	mean, terminal-regions
\ctmh	$\hat{c}_t^{[m]}$	mean, terminal-regions with hat
\ctmt	$\tilde{c}_t^{[m]}$	mean, terminal-regions
\fxk	$f_k(x)$	f_k(x)
\Lp	L'	
\Ldp	L''	
\Lpleft	L'_left	
\Lxyim	$L\left(y^{(i)}, f^{[m-1]}(\mathbf{x}^{(i)})\right)$	

ml-feature-sel.tex

Macro	Notation	Comment
\xjNull	x_{j_0}	
\xjEins	x_{j_1}	
\xl	$\mathbf{x}_l$	
\pushcode		IGNORE_NOTATION

ml-gp.tex

Macro	Notation	Comment
\gp	$\mathcal{GP}\left(m(\boldsymbol{x}), k\left(\boldsymbol{x}, \boldsymbol{x}^{\prime}\right)\right)$	Gaussian Process Definition
\mvec	$\mathbf{m}$	Gaussian process mean vector
\Kmat	$\mathbf{K}$	estimated base learner
\ls	ℓ	length-scale

ml-interpretable.tex

Macro	Notation	Comment
`\fj`	f_j	marginal function f_j, depending on feature j
`\fS`	f_S	marginal function f_S depending on feature set S
`\fC`	f_C	marginal function f_C depending on feature set C
`\fhj`	$\hat{f}_j$	marginal function fh_j, depending on feature j
`\fhS`	$\hat{f}_S$	marginal function fh_S depending on feature set S
`\fhC`	$\hat{f}_C$	marginal function fh_C depending on feature set C
`\XSmat`	$\mathbf{X}_S$	Design matrix subset
`\XCmat`	$\mathbf{X}_C$	Design matrix subset
`\Scupj`	$S \cup \{j\}$	coalition S but without player j
`\Scupk`	$S \cup \{k\}$	coalition S but without player k
`\SsubP`	$S \subseteq P$	coalition S subset of P
`\SsubPnoj`	$S \subseteq P \setminus \{j\}$	coalition S subset of P without player j
`\SsubPnojk`	$S \subseteq P \setminus \{j, k\}$	coalition S subset of P without player k
`\phiij`	$\hat{\phi}_j^{(i)}$	Shapley value for feature j and observation i
`\Gspace`	$\mathcal{G}$	Hypothesis space for surrogate model
`\neigh`	$\phi_{\mathbf{x}}$	Proximity measure
`\zv`	$\mathbf{z}$	Sampled datapoints for surrogate
`\Zspace`	$\mathcal{Z}$	Space of sampled datapoints
`\Gower`	d_G	Gower distance

ml-lm.tex

Macro	Notation	Comment
`\thx`	$\boldsymbol{\theta}^T\mathbf{x}$	linear score: theta^T x

ml-mbo.tex

Macro	Notation	Comment
`\sxh`	$\hat{s}(x)$	uncertainty shat(x)
`\vxh`	$\hat{s}^2(x)$	squared uncertanty
`\matK`	$\mathbf{K}$	??
`\kstarx`	$\mathbf{k}_*(x)$	??
`\xpi`	$x^{*(\#1)}$	??
`\vhx`	$\hat{s}^2(\mathbf{x})$	local estimated variance at point x
`\shx`	$\hat{s}(\mathbf{x})$	local estimated uncertainty at point x
`\sh`	$\hat{s}$	local estimated uncertainty
`\px`	$\boldsymbol{x}^*$	??
`\equote`	"#1″	??
`\vecx`	$\boldsymbol{x}$	DO NOT USE THIS
`\yx`	$y(\boldsymbol{x})$	??
`\X`	$\mathcal{X}$	domain / search space
`\yv`	$\boldsymbol{y}$	
`\fhx`	$\hat{f}(\mathbf{x})$	surrogate (x), better use \mhx for predicted value
`\minit`	m_{init}	Size of the initial design
`\lambdai`	$\boldsymbol{\lambda}^{[i]}$	input for black box optimization
`\lambdanew`	$\boldsymbol{\lambda}^{\text{new}}$	new proposed configuration
`\metadata`	$\left\{\left(\lambda^{[i]}, \Psi^{[i]}\right)\right\}$	Metadata for the Gaussian process
`\lambdavec`	$\boldsymbol{\lambda}^{[1]}, ..., \boldsymbol{\lambda}^{[m_{\text{init}}]}$	Vector of different inputs
`\lambdab`	$\boldsymbol{\lambda}$	input
`\lambdaopt`	$\boldsymbol{\lambda}^*$	Minimum of the black box function Psi

ml-nn.tex

Macro	Notation	Comment
\neurons	$z_1, \dots, z_M$	vector of neurons
\hidz	$\mathbf{z}$	vector of hidden activations
\biasb	$\mathbf{b}$	bias vector
\biasc	c	bias in output
\wtw	$\mathbf{w}$	weight vector (general)
\Wmat	$\mathbf{W}$	weight vector (general)
\wtu	$\mathbf{u}$	weight vector of output neuron
\Oreg	$R_{reg}(\theta \| X, y)$	regularized objective function
\Ounreg	$R_{emp}(\theta \| X, y)$	unconstrained objective function
\Pen	$\Omega(\theta)$	penalty
\Oregweight	$R_{reg}(w \| X, y)$	regularized objective function with weight
\Oweight	$R_{emp}(w \| X, y)$	unconstrained objective function with weight
\Oweighti	$R_{emp}(w_i \| X, y)$	unconstrained objective function with weight w_i
\Oweightopt	$J(w^* \| X, y)$	unconstrained objective function withoptimal weight
\Oopt	$\hat{J}(\theta \| X, y)$	optimal objective function
\Odropout	$J(\theta, \mu \| X, y)$	dropout objective function
\Loss	$L(y, f(\mathbf{x}, \boldsymbol{\theta}))$	
\Lmomentumnest	$L(y^{(i)}, f(x^{(i)}, \boldsymbol{\theta} + \varphi \boldsymbol{\nu}))$	momentum risk
\Lmomentumtilde	$L(y^{(i)}, f(x^{(i)}, \tilde{\boldsymbol{\theta}}))$	Nesterov momentum risk
\Lmomentum	$L(y^{(i)}, f(x^{(i)}, \boldsymbol{\theta}))$	
\Hess	$\mathbf{H}$	
\nub	$\boldsymbol{\nu}$	
\uauto	$L(x, g(f(x)))$	undercomplete autoencoder objective function
\dauto	$L(x, g(f(\tilde{x})))$	denoising autoencoder objective function
\deltab	$\boldsymbol{\delta}$	
\Lossdeltai	$L(y^{(i)}, f(\mathbf{x}^{(i)} + \boldsymbol{\delta} \| \boldsymbol{\theta}))$	
\Lossdelta	$L(y, f(\mathbf{x} + \boldsymbol{\delta} \| \boldsymbol{\theta}))$	

ml-rf.tex

Macro	Notation	Comment
`\betam`	$\beta^{[m]}$	baselearner with argument for m
`\betai`	$\beta^{[1]}$	baselearner with argument for 1
`\betaM`	$\beta^{[M]}$	baselearner with argument for M

ml-survival.tex

Macro	Notation	Comment
`\Ti`	$T^{(\#1)}$	??
`\Ci`	$C^{(\#1)}$	??
`\oi`	$o^{(\#1)}$	??
`\ti`	$t^{(\#1)}$	??
`\lambdai`	$\lambda^{(\#1)}$	already defined differenlty in ml-bayesopt
`\deltai`	$\delta^{(\#1)}$	
`\Lxdi`	$L\left(\boldsymbol{\delta}, f(\mathbf{x})\right)$	

ml-svm.tex

Macro	Notation	Comment
`\sv`	SV	supportvectors
`\HS`	Φ	H, hilbertspace
`\sl`	ζ	
`\slvec`	$\left(\zeta^{(1)}, \zeta^{(n)}\right)$	slack variables (SVM)
`\sli`	$\zeta^{(i)}$	slack variable (SVM)
`\alphah`	$\hat{\alpha}$	alpha-hat
`\alphav`	$\boldsymbol{\alpha}$	vector alpha (bold)
`\alphavh`	$\hat{\boldsymbol{\alpha}}$	vector alpha-hat
`\phix`	$\phi(\mathbf{x})$	\phi(x)
`\phixt`	$\phi(\tilde{\mathbf{x}})$	\phi(x-tilde)

ml-trees.tex

Macro	Notation	Comment
`\Np`	$\mathcal{N}$	(Parent) node N
`\Npk`	$\mathcal{N}_k$	node N_k
`\Nl`	$\mathcal{N}_1$	Left node N_1
`\Nr`	$\mathcal{N}_2$	Right node N_2
`\pikN`	$\pi_k^{(\mathcal{N})}$	class probability node N
`\pikNh`	$\hat{\pi}_{\#1}^{(\mathcal{N})}$	estimated class probability node N
`\pikNlh`	$\hat{\pi}_{\#1}^{(\mathcal{N}_1)}$	
`\pikNrh`	$\hat{\pi}_{\#1}^{(\mathcal{N}_2)}$	

probmodel.tex

Macro	Notation	Comment
`\muk`	$\boldsymbol{\mu}_{\boldsymbol{k}}$	