1228 Probabilistic Graphical Models
H. Zhu and J. Boedecker
April 2, 2026

Basic Notation

We use $\mathbf{R}$ to denote the set of real numbers, $\mathbf{R}_+$ to denote the set of nonnegative real numbers, and $\mathbf{R}_{++}$ to denote the set of positive real numbers. the set of real n-vectors is denoted $\mathbf{R}^n$, and the set of real $m \times n$ matrices is denoted $\mathbf{R}^{m\times n}$. The symbol $\mathbf{1}$ denotes a vector all of whose components are one (with dimension determined from context).

The curled inequality symbols $\preceq$ and $\prec$ denote componentwise inequalities between vectors: $x \preceq y$ means that $x_i \le y_i$ for all i, and $x \prec y$ means that $x_i < y_i$ for all i. The inequality symbols $\succ$ and $\succeq$ are defined similarly, but with the direction of the inequalities reversed.

We use the notation $f\colon \mathbf{R}^p \to \mathbf{R}^q$ to mean that f is an $\mathbf{R}^q$ valued function on some subset of $\mathbf{R}^p$, specifically, its domain, which we denote $\mathbf{dom}\, f$. We can think of our use of the notation $f\colon \mathbf{R}^p \to \mathbf{R}^q$ as a declaration of the function *type*, as in a computer language: $f\colon \mathbf{R}^p \to \mathbf{R}^q$ means that the function f takes as argument a real p-vector, and returns a real q-vector. The set $\mathbf{dom}\, f$, the domain of the function f, specifies the subset of $\mathbf{R}^p$ of points x for which $f(x)$ is defined. As an example, we describe the logarithm function as $\log\colon \mathbf{R} \to \mathbf{R}$, with $\mathbf{dom} \log = \mathbf{R}_{++}$. The notation $\log\colon \mathbf{R} \to \mathbf{R}$ means that the logarithm function accepts and returns a real number; $\mathbf{dom} \log = \mathbf{R}_{++}$ means that the logarithm is defined only for positive numbers.

Sets, vectors, and matrices

$\mathbf{R}$	Real numbers.
$\mathbf{R}^n$	Real n-vectors ($n \times 1$ matrices).
$\mathbf{R}^{m\times n}$	Real $m \times n$ matrices.
$\mathbf{R}_+, \mathbf{R}_{++}$	Nonnegative, positive real numbers.
$\mathbf{Z}$	Integers.
$\mathbf{Z}_+, \mathbf{Z}_{++}$	Nonnegative, positive integers.
$\lvert C\rvert$	Cardinality of set C.
I_C	Indicator function of set C.
$\mathbf{1}$	Vector with all components one.
e_i	ith standard basis vector.
$x \odot y$	Componentwise multiplication of vectors x and y.

I	Identity matrix.
$X_{i:}, X_{:i}$	The ith row/column of matrix X, represented as a column vector.
X^T	Transpose of matrix X.
X^k	(Square) matrix X to the kth power.
$\mathbf{tr}\, X$	Trace of matrix X.
$\mathbf{diag}(x)$	Diagonal matrix with diagonal entries $x_1, \ldots, x_n$.
$\mathbf{rank}\, A$	Rank of matrix A.

Functions and derivatives

$f \colon A \to B$	f is a function on the set $\mathbf{dom}\, f \subseteq A$ into the set B.
$\mathbf{dom}\, f$	Domain of function f.
∇f	Gradient of function f.
$\nabla^2 f$	Hessian of function f.

Norms and distances

$\|\cdot\|$	A norm.
$\|x\|_1$	l_1-norm of vector x.
$\|x\|_2$	Euclidean (or l_2-) norm of vector x.
$\|x\|_\infty$	l_∞-norm of vector x.
$\mathbf{dist}(A, B)$	Distance between sets (or points) A and B.

Generalized inequalities

$x \preceq y$	Componentwise inequality between vectors x and y.
$x \prec y$	Strict componentwise inequality between vectors x and y.

Probability

$\mathbf{P}(S)$	Probability of event S.
$(X \perp\!\!\!\perp Y \mid Z)$	Conditional independence of random variables X and Y given Z.
$\mathbf{E}\, X$	Expected value of random variable X.
$\mathbf{var}\, X$	Variance of random variable X.
$\sigma(X)$	Standard deviation of random variable X.

$\mathbf{cov}(X, Y)$	Covariance of random variables X and Y.
$\rho(X, Y)$	Correlation coefficient of random variables X and Y.
$r(X, Y)$	Regression coefficient of random variables X and Y.
$p(x)$	Density function of continuous random variable X.
$F(x)$	Cumulative distribution function of continuous random variable X.
$l_x(\theta)$	Log-likelihood function of θ given the observation $X = x$.
$\mathbf{supp}(p)$	Support of density function p.
$\mathcal{N}(\mu, \sigma^2)$	Gaussian distribution with mean μ, variance σ^2.
$\Phi(x)$	Cumulative distribution function of $\mathcal{N}(0, 1)$ random variable X.
$\mathcal{U}(x, y)$	Uniform distribution on interval $[x, y]$.
$\mathrm{Exp}(\lambda)$	Exponential distribution with parameter λ.
$\mathrm{Beta}(\alpha, \beta)$	Beta distribution with shape parameters α and β.
$\mathrm{Dir}(\alpha)$	Dirichlet distribution with concentration parameters $\alpha = (\alpha_1, \ldots, \alpha_k)$.

Graph

$\mathbf{pa}(V)$	Parents of node V.
$\mathbf{mb}(V)$	Markov blanket of node V.
$\mathbf{adj}(V)$	Adjacent nodes of node V.