NOTATIONAL CONVENTIONS

Vectors and matrices. By default, all vectors are column ones; to write them down, we use "Matlab notation": $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ is written as $[1; 2; 3]$. More generally, for vectors/matrices $A, B, ..., Z$ of the same "width" (or vectors/matrices $A, B, C, ..., Z$ of the same "height"), $[A; B; C; ...; D]$ is the matrix obtained by vertical (or horizontal) concatenation of A, B, C, etc. Examples: For what in the "normal" notation is written down as $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 5 & 6 \end{bmatrix}$, $C = \begin{bmatrix} 7 \\ 8 \end{bmatrix}$, we have

$$[A; B] = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = [1, 2; 3, 4; 5, 6], \ [A, C] = \begin{bmatrix} 1 & 2 & 7 \\ 3 & 4 & 8 \end{bmatrix} = [1, 2, 7; 3, 4, 8].$$

Blanks in matrices replace (blocks of) zero entries. For example,

$$\begin{bmatrix} 1 & & \\ 2 & & \\ 3 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 3 & 4 & 5 \end{bmatrix}.$$

$\mathrm{Diag}\{A_1, A_2, ..., A_k\}$ stands for a block-diagonal matrix with diagonal blocks A_1, A_2, ..., A_k. For example,

$$\mathrm{Diag}\{1, 2, 3\} = \begin{bmatrix} 1 & & \\ & 2 & \\ & & 3 \end{bmatrix}, \ \mathrm{Diag}\{[1, 2]; [3; 4]\} = \begin{bmatrix} 1 & 2 & \\ & & 3 \\ & & 4 \end{bmatrix}.$$

For an $m \times n$ matrix A, $\mathrm{dg}(A)$ is the diagonal of A—a vector of dimension $\min[m, n]$ with entries A_{ii}, $1 \le i \le \min[m, n]$.

Standard linear spaces in our book are $\mathbf{R}^n$ (the space of n-dimensional column vectors), $\mathbf{R}^{m \times n}$ (the space of $m \times n$ real matrices), and $\mathbf{S}^n$ (the space of $n \times n$ real symmetric matrices). All these linear spaces are equipped with the standard inner product:

$$\langle A, B \rangle = \sum_{i,j} A_{ij} B_{ij} = \mathrm{Tr}(AB^T) = \mathrm{Tr}(BA^T) = \mathrm{Tr}(A^T B) = \mathrm{Tr}(B^T A);$$

in the case when $A = a$ and $B = b$ are column vectors, this simplifies to $\langle a, b \rangle = a^T b = b^T a$, and when A, B are symmetric, there is no need to write B^T in $\mathrm{Tr}(AB^T)$.

Usually, we denote vectors by lowercase, and matrices by uppercase letters; sometimes, however, lowercase letters are used also for matrices.

Given a linear mapping $\mathcal{A}(x) : E_x \to E_y$, where E_x, E_y are standard linear spaces, one can define the *conjugate* mapping $\mathcal{A}^*(y) : E_y \to E_x$ via the identity

$$\langle \mathcal{A}(x), y \rangle = \langle x, \mathcal{A}^*(y) \rangle \ \forall (x \in E_x, y \in E_y).$$

One always has $(\mathcal{A}^*)^* = \mathcal{A}$. When $E_x = \mathbf{R}^n$, $E_y = \mathbf{R}^m$ and $\mathcal{A}(x) = Ax$, one has $\mathcal{A}^*(y) = A^T y$; when $E_x = \mathbf{R}^n$, $E_y = \mathbf{S}^m$, so that $\mathcal{A}(x) = \sum_{i=1}^n x_i A_i$, $A_i \in \mathbf{S}^m$, we have

$$\mathcal{A}^*(Y) = [\mathrm{Tr}(A_1Y); ...; \mathrm{Tr}(A_nY)].$$

$\mathbf{Z}^n$ is the set of n-dimensional integer vectors.

Norms. For $1 \leq p \leq \infty$ and for a vector $x = [x_1; ...; x_n] \in \mathbf{R}^n$, $\|x\|_p$ is the standard p-norm of x:

$$\|x\|_p = \left\{ \begin{array}{ll} \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}, & 1 \leq p < \infty, \\ \max_i |x_i| = \lim_{p' \to \infty} \|x\|_{p'}, & p = \infty. \end{array} \right.$$

The spectral norm (the largest singular value) of a matrix A is denoted by $\|A\|_{2,2}$; notation for other norms of matrices is specified when used.

Standard cones. $\mathbf{R}_+$ is the nonnegative ray on the real axis; $\mathbf{R}^n_+$ stands for the n-*dimensional nonnegative orthant*, the cone comprised of all entrywise nonnegative vectors from $\mathbf{R}^n$; $\mathbf{S}^n_+$ stands for the *positive semidefinite cone in* $\mathbf{S}^n$, the cone comprised of all positive semidefinite matrices from $\mathbf{S}^n$.

Miscellaneous.
• For matrices A, B, relation $A \preceq B$, or, equivalently, $B \succeq A$, means that A, B are symmetric matrices of the same size such that $B - A$ is positive semidefinite; we write $A \succeq 0$ to express the fact that A is a symmetric positive semidefinite matrix. Strict version $A \succ B$ ($\Leftrightarrow B \prec A$) of $A \succeq B$ means that $A - B$ is positive definite (and, as above, A and B are symmetric matrices of the same size).
• Linear Matrix Inequality (LMI, a.k.a. *semidefinite constraint*) in variables x is the constraint on x stating that a symmetric matrix affinely depending on x is positive semidefinite. When $x \in \mathbf{R}^n$, LMI reads

$$A_0 + \sum_i x_i A_i \succeq 0 \qquad [A_i \in \mathbf{S}^m, 0 \leq i \leq n].$$

• $\mathcal{N}(\mu, \Theta)$ stands for the Gaussian distribution with mean μ and covariance matrix Θ. Poisson(μ) denotes Poisson distribution with parameter $\mu \in \mathbf{R}_+$, i.e., the distribution of a random variable taking values $i = 0, 1, 2, ...$ with probabilities $\frac{\mu^i}{i!}\mathrm{e}^{-\mu}$. Uniform($[a, b]$) is the uniform distribution on segment $[a, b]$.
• For a probability distribution P,

- $\xi \sim P$ means that ξ is a random variable with distribution P. Sometimes we express the same fact by writing $\xi \sim p(\cdot)$, where p is the density of P taken w.r.t. some reference measure (the latter always is fixed by the context);
- $\mathbf{E}_{\xi\sim P}\{f(\xi)\}$ is the expectation of $f(\xi)$, $\xi \sim P$; when P is clear from the context, this notation can be shortened to $\mathbf{E}_\xi\{f(\xi)\}$, or $\mathbf{E}_P\{f(\xi)\}$, or even $\mathbf{E}\{f(\xi)\}$. Similarly, $\mathrm{Prob}_{\xi\sim P}\{...\}$, $\mathrm{Prob}_\xi\{...\}$, $\mathrm{Prob}_P\{...\}$, and $\mathrm{Prob}\{...\}$ denote the P-probability of the event specified inside the braces.

• $O(1)$'s stand for positive *absolute* constants—positive reals with numerical values (completely independent of the parameters of the situation at hand) which we do not want or are too lazy to write down explicitly, as in $\sin(x) \leq O(1)|x|$.
• $\int_\Omega f(\xi)\Pi(d\xi)$ stands for the integral, taken w.r.t. measure Π over domain Ω, of function f.