Building blocks for this lecture

Ryota Tomioka

August 14, 2013

1 Equalities

1. Euclidean norm
$$\|\boldsymbol{x}\|_2^2 = \boldsymbol{x}^\top \boldsymbol{x}.$$

2. Expanding the squared norm of the sum of two vectors
$$\|\boldsymbol{x} + \boldsymbol{y}\|_2^2 = \|\boldsymbol{x}\|_2^2 + 2\boldsymbol{x}^\top \boldsymbol{y} + \|\boldsymbol{y}\|_2^2.$$

3. Singular value decomposition
$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^\top,$$
where $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthogonal matrices and $\boldsymbol{S}$ is diagonal.

4. Orthogonal matrix
$$\boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^\top = \boldsymbol{I}.$$

5. Trace identity
$$\mathrm{Tr}(\boldsymbol{A}\boldsymbol{B}) = \mathrm{Tr}(\boldsymbol{B}\boldsymbol{A}).$$

2 Inequalities

6. Any norm $\|\cdot\|$ satisfies
   (a) Positive homogeneity
   $$\|a\boldsymbol{x}\| = |a|\|\boldsymbol{x}\|.$$
   (b) Triangular inequality
   $$\|\boldsymbol{x}\| - \|\boldsymbol{y}\| \leq \|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|.$$
   (c) Zero means zero
   $$\|\boldsymbol{x}\| = 0 \quad \Rightarrow \quad \boldsymbol{x} = \boldsymbol{0}.$$

(d) Positivity

$$\|\boldsymbol{x}\| \geq 0.$$

(implied by the positive homogeneity and the triangular inequality).

7. Euclidean norm is dual to itself

$$\boldsymbol{x}^\top \boldsymbol{y} \leq \|\boldsymbol{x}\|_2 \|\boldsymbol{y}\|_2.$$

8. ℓ_∞- and ℓ_1-norms are dual to each other

$$\boldsymbol{x}^\top \boldsymbol{y} \leq \|\boldsymbol{x}\|_\infty \|\boldsymbol{y}\|_1,$$

where $\|\boldsymbol{x}\|_\infty = \max_j |x_j|$ and $\|\boldsymbol{y}\|_1 = \sum_{j=1}^p |y_j|$.

9. Compatibility between ℓ_1 and ℓ_2 (Euclidean norm): if $\boldsymbol{x}$ is k-sparse (only k non-zero entries), then

$$\|\boldsymbol{x}\|_1 \leq \sqrt{k} \|\boldsymbol{x}\|_2.$$

10. Euclidean norm increases when adding two orthogonal vectors: if $\boldsymbol{x}^\top \boldsymbol{y} = 0$,

$$\|\boldsymbol{x}\|_2 \leq \|\boldsymbol{x} + \boldsymbol{y}\|_2 \quad \text{and} \quad \|\boldsymbol{y}\|_2 \leq \|\boldsymbol{x} + \boldsymbol{y}\|_2$$

3 Probability theory

11. Linearity of expectation

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y].$$

12. Reproducibility of Gaussian: if $X, Y \sim \mathcal{N}(0, \sigma^2)$,

$$aX + bY \sim \mathcal{N}(0, (a^2 + b^2)\sigma^2).$$

13. Gaussian covariance: if $\boldsymbol{x} \sim \mathcal{N}(0, \boldsymbol{\Sigma})$,

$$\mathbb{E}[\boldsymbol{x}\boldsymbol{x}^\top] = \boldsymbol{\Sigma}.$$

14. Union bound: the probability of any of the n events $E_1, \ldots, E_n$ being true is bounded by the sum of the probabilities:

$$P(\cup_{i=1}^n E_i) \leq \sum_{i=1}^n P(E_i).$$

15. Max of Gaussians: let $z_j \sim \mathcal{N}(0, \sigma_j^2)$ $(j = 1, \ldots, p)$, then

$$\Pr\left(\max_j |z_j| \geq 2R\sqrt{\log p}\right) \leq \frac{2}{p} \quad \text{where} \quad R := \max_j \sigma_j.$$

In other words, with probability at least $1 - 2/p$, we have

$$\max_j |z_j| \leq 2R\sqrt{\log p}.$$