CS205 Review Session #6 Notes

More Norms

Let $\mathbf{A}$ be an $n \times n$ positive definite matrix. We can write $\mathbf{A}$ as $\mathbf{A} = \mathbf{M}\mathbf{M}^T$ where $\mathbf{M}$ is an appropriate $n \times n$ matrix. There are many choices for $\mathbf{M}$. For example, using Cholesky factorization, we can write $\mathbf{A} = \mathbf{L}\mathbf{L}^T$ where $\mathbf{L}$ is a lower triangular matrix. Alternatively, using the diagonal form of $\mathbf{A}$ we can write:

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T = \mathbf{Q}\mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{1/2}\mathbf{Q} = (\mathbf{Q}\mathbf{\Lambda}^{1/2})(\mathbf{Q}\mathbf{\Lambda}^{1/2})^T$$

Using any such matrix $\mathbf{M}$ allows us to express

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{A}} = \mathbf{u}^T\mathbf{A}\mathbf{v} = \mathbf{u}^T\mathbf{M}\mathbf{M}^T\mathbf{v} = \mathbf{M}^T\mathbf{u} \cdot \mathbf{M}^T\mathbf{v}$$

Therefore, the inner product induced by $\mathbf{A}$ is equivalent to transforming our vector space into a new vector space via the mapping $\mathbf{x} \mapsto \mathbf{M}^T\mathbf{x}$, and then taking the usual Euclidean dot product into the transformed space.

This can be used to prove that the norm $\|\mathbf{x}\|_{\mathbf{A}} = \sqrt{\mathbf{x}\mathbf{A}^T\mathbf{x}} = \|\mathbf{M}^T\mathbf{x}\|_2$ satisfies the properties of a norm. In a similar way, we can show that the inner product induced by $\mathbf{A}$ has the properties of a regular dot product.

Conjugate Vectors

Let $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ be a set of vectors in $\mathbb{R}^n$. The Gram-Schmidt algorithm for creating an orthogonal set $\{\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \ldots, \tilde{\mathbf{x}}_k\}$ is given by the recurrence:

$$\tilde{\mathbf{x}}_i = \mathbf{x}_i - \sum_{j=1}^{i-1} \frac{\mathbf{x}_i \cdot \tilde{\mathbf{x}}_j}{\tilde{\mathbf{x}}_j \cdot \tilde{\mathbf{x}}_j} \tilde{\mathbf{x}}_j$$

The corresponding algorithm for creating a set of *A-orthogonal* vectors is

$$\tilde{\mathbf{x}}_i = \mathbf{x}_i - \sum_{j=1}^{i-1} \frac{\mathbf{x}_i \cdot \mathbf{A}\tilde{\mathbf{x}}_j}{\tilde{\mathbf{x}}_j \cdot \mathbf{A}\tilde{\mathbf{x}}_j} \tilde{\mathbf{x}}_j$$

Note that in the computation of $\tilde{\mathbf{x}}_i$ we subtract a linear combination of $\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \ldots, \tilde{\mathbf{x}}_{i-1}$ from $\mathbf{x}_i$.

Problem 3.5

For an $n \times n$ matrix $\mathbf{A}$ we have

$$|\mathbf{x}^T\mathbf{A}\mathbf{x}| = |\mathbf{x} \cdot \mathbf{A}\mathbf{x}| \leq \|\mathbf{x}\|_2 \|\mathbf{A}\mathbf{x}\|_2 \leq \|\mathbf{x}\|_2 \|\mathbf{A}\|_2 \|\mathbf{x}\|_2 = \sigma_{\max}^{\mathbf{A}} \|\mathbf{x}\|_2^2$$

Furthermore, if $\mathbf{A}$ is *symmetric* we have $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$ and

$$\frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \frac{\mathbf{x}^T\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T\mathbf{x}}{\mathbf{x}^T\mathbf{Q}\mathbf{Q}^T\mathbf{x}} \stackrel{\mathbf{y}=\mathbf{Q}^T\mathbf{x}}{=} \frac{\mathbf{y}^T\mathbf{\Lambda}\mathbf{y}}{\mathbf{y}^T\mathbf{y}} = \frac{\sum \lambda_i y_i^2}{\sum y_i^2}$$

This allows us to observe that:

$$\lambda_{\min} \leq \frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \leq \lambda_{\max}$$