**Matrices and vectors:** All these are very important facts that we will use repeatedly, and should be internalized.

- *Inner product and outer products.* Let $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ be vectors in $\mathbb{R}^n$. Then $u^T v = \langle u, v \rangle = \sum_{i=1}^n u_i v_i$ is the inner product of $u$ and $v$.

  The outer product $uv^T$ is an $n \times n$ rank 1 matrix $B$ with entries $B_{ij} = u_i v_j$. The matrix $B$ is a very useful operator. Suppose $v$ is a unit vector. Then, $B$ sends $v$ to $u$ i.e. $Bv = uv^T v = u$, but $Bw = \mathbf{0}$ for all $w \in v^\perp$.

- *Matrix Product.* For any two matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times p}$, the standard matrix product $C = AB$ is the $m \times p$ matrix with entries $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$. Here are two very useful ways to view this.

  Inner products: Let $r_i$ be the $i$-th row of $A$, or equivalently the $i$-th column of $A^T$, the transpose of $A$. Let $b_j$ denote the $j$-th column of $B$. Then $c_{ij} = r_i^T c_j$ is the dot product of $i$-column of $A^T$ and $j$-th column of $B$.

  Sums of outer products: $C$ can also be expressed as outer products of columns of $A$ and rows of $B$.

  *Exercise:* Show that $C = \sum_{k=1}^n a_k b_k^T$ where $a_k$ is the $k$-th column of $A$ and $b_k$ is the $k$-th row of $B$ (or equivalently the $k$-column of $B^T$).

- *Trace inner product of matrices.* For any $n \times n$ matrix $A$, the trace is defined as the sum of diagonal entries, $Tr(A) = \sum_i a_{ii}$. For any two $m \times n$ matrices $A$ and $B$ one can define the Frobenius or Trace inner product $\langle A, B \rangle = \sum_{ij} a_{ij} b_{ij}$. This is also denoted as $A \bullet B$.

  *Exercise:* Show that $\langle A, B \rangle = Tr(A^T B) = Tr(BA^T)$.

- *Bilinear forms.* For any $m \times n$ matrix $A$ and vectors $u \in \mathbb{R}^m, v \in \mathbb{R}^n$, the product $u^T Av = \langle u, Av \rangle = \sum_{i=1}^n \sum_{j=1}^n a_{ij} u_i v_j$.

  *Exercise:* Show that $u^T Av = \langle uv^T, A \rangle$. This relates bilinear product to matrix inner product.

- *PSD matrices.* Let $A$ be a symmetric $n \times n$ matrix with entries $a_{ij}$. Recall that we defined $A$ to be PSD if there exist vectors $v_i$ for $i = 1, \ldots, n$ (in some arbitrary dimensional space) such that $a_{ij} = v_i \cdot v_j$ for each $1 \le i, j \le n$.

  The following properties are all equivalent ways to characterizing PSD matrices:

  1. $a_{ij} = v_i \cdot v_j$ for each $1 \le i, j \le n$. $A$ is called the Gram matrix of vectors $v_i$. So if $V$ is the matrix obtained by stacking these vectors with $i$-th column $v_i$, then $A = V^T V$.
  2. $A$ is symmetric and $x^T Ax \ge 0$ for all $x \in \mathbb{R}^n$.
  3. $A$ is symmetric and has all eigenvalues non-negative.
  4. $A = \sum_{i=1}^n \lambda_i u_i u_i^T$ for $\lambda_i \ge 0$ and $u_i$ are an orthonormal set of vectors. The $\lambda_i$ are the eigenvalues of $A$ and $u_i$ are the corresponding eigenvectors.

  *Exercise:* Show that (1) → (2) and (2) → (3).

  **Solution:**

**(1) → (2):** *Note that* (1) *implies that* $A = V^TV$, *where* $V = (v_1, \dots, v_n)$. *Then for any* $x \in \mathbb{R}^n$, *we have that*

$$x^TAx = x^TV^TVx = \|Vx\|^2 \geq 0,$$

*where* $\|Vx\|^2$ *is the squared Euclidean norm.*

**(2) → (3):** *Since* $A$ *is symmetric,* $A$ *is diagonizable and only real eigen values. Now assume that* $A$ *has a negative eigen value* $\lambda < 0$ *with corresponding eigen vector* $v \neq 0$ *such that* $Av = \lambda v$. *But then,* $v^TAv = \lambda\|v\|^2 < 0$, *contradicting* (2).

(3) → (4) follows from the well-known spectral theorem (that we do not prove here) that any symmetric matrix $B$ has real eigenvalues and its eigenvectors are orthogonal. That is, $B = \sum_i \beta_i u_i u_i^T$ where $\beta_i \in \mathbb{R}$ and $u_i$ is an orthonormal set of vectors.

*Exercise:* Show that (4) → (1) using the two views of matrix products discussed above. **Solution:** *Let* $U = (\sqrt{\lambda_1}u_1, \dots, \sqrt{\lambda_n}u_n)$ *(i.e. with the indicated columns). Then*

$$A = \sum_{i=1}^{n} \lambda_i u_i u_i^T = \sum_{i=1}^{n} (\sqrt{\lambda_i}u_i)(\sqrt{\lambda_i}u_i)^T = UU^T \ .$$

*Now letting* $(v_1, \dots, v_n)^T = U$ *denote the columns of* $U$, *we directly get that* $A_{ij} = (UU^T)_{ij} = v_i \cdot v_j$.

*Exercise:* If $A$ and $B$ are $n \times n$ PSD matrics. Show that $A + B$ is also PSD. **Solution:** *Note that* $x^TAx \geq 0$ *and* $x^TBx \geq 0$, $\forall x \in \mathbb{R}^n$, *clearly implies that* $x^T(A+B)x = x^TAx + x^TBx \geq 0$ $\forall x \in \mathbb{R}^n$. *Thus, by (2)* $A + B$ *is also PSD.*

Hint: It is easiest to use definition (2) of PSD above.

*Exercise:* Show the above using (1) instead of (2). In particular if $a_{ij} = v_i \cdot v_j$ and $b_{ij} = w_i \cdot w_j$ can you construct vectors $y_i$ using these $v_i$ and $w_i$ such that $a_{ij} + b_{ij} = y_i \cdot y_j$?

**Solution:** *Define* $z_1, \dots, z_n$ *by the relation* $z_i^T = (v_i^T, w_i^T)$ *for* $i \in [n]$. *Then clearly* $\langle z_i, z_j \rangle = \langle v_i, v_j \rangle + \langle w_i, w_j \rangle = a_{ij} + b_{ij}$.

- *Tensors.* Let $v \in R^n$. We define the two-fold tensor $v^{\otimes 2}$ as the $n \times n$ matrix with $(i,j)$-th entry $v_i \cdot v_j$. This is same as $vv^T$, but it is useful to view $v^{\otimes 2}$ as an $n^2$ dimensional vector. Similarly, if $v \in R^n$ and $w \in R^m$, $v \otimes w = vw^T$ is viewed as an $nm$ dimensional vector.

  *Exercise:* Show that if $v, w \in R^n$ and $x, y \in R^m$, then $\langle v \otimes x, w \otimes y \rangle = \langle v, w \rangle \langle x, y \rangle$. One can remember this rule as, the dot product of tensors is the product of their vector dot products. **Solution:**

$$\begin{aligned} \langle v \otimes x, w \otimes y \rangle &= \sum_{i \in [n], j \in [m]} (v \otimes x)_{ij} (w \otimes y)_{ij} = \sum_{i \in [n], j \in [m]} v_i x_j w_i y_j \\ &= (\sum_{i \in [n]} v_i w_i)(\sum_{j \in [m]} x_j y_j) = \langle v, w \rangle \langle x, y \rangle \ . \end{aligned}$$

  Similarly, one can generalize this to higher order tensors. For now we just discuss the $k$-fold tensor of a vector by itself. If $v \in R^n$ $v^{\otimes k}$ is the $n^k$ dimensional vector with the $(i_1, \dots, i_k)$ entry equal to the product $v_{i_1} v_{i_2} \cdots v_{i_k}$.

  *Exercise:* Show (by just expanding things out) that if $v, w \in \mathbb{R}^n$ then $v^{\otimes k}, w^{\otimes k} = (\langle v, w \rangle)^k$.

**Solution:**

$$\begin{aligned}\langle v^{\otimes k}, w^{\otimes k}\rangle &= \sum_{i_1,\dots,i_k\in[n]} v^{\otimes k}_{i_1 i_2\dots i_k} w^{\otimes k}_{i_1 i_2\dots i_k} = \sum_{i_1,\dots,i_k\in[n]} (v_{i_1}\cdots v_{i_k})(w_{i_1}\cdots w_{i_k}) \\ &= \sum_{i_1,\dots,i_k\in[n]} (v_{i_1}w_{i_1})\cdots(v_{i_k}w_{i_k}) = (\sum_{i_1\in[n]} v_{i_1}w_{i_1})\cdots(\sum_{i_k\in[n]} v_{i_k}w_{i_k}) = \langle v,w\rangle^k \ ,\end{aligned}$$

*as needed.*

*Exercise:* Let $p(x)$ a univariate polynomial with non-negative coefficients. Let $A$ be a $n\times n$ PSD matrix with entries $a_{ij}$, and let $p(A)$ denote the matrix which has its $(i,j)$-entry $p(a_{ij})$. Show that $p(A)$ is also PSD.

Hint: Use that $a_{ij} = \langle v_i, v_j\rangle$ for each $i,j$, and construct suitable vectors $v_i'$ and $v_j'$ such that $p(a_{ij}) = v_i'\cdot v_j'$. Use the property $\langle v^{\otimes k}, w^{\otimes k}\rangle = (\langle v,w\rangle)^k$ of dot products tensors stated above.

**Solution:** *Since $A$ is PSD we can write $A_{ij} = \langle v_i, v_j\rangle$ for vectors $v_1,\dots,v_n$. Let $p(x) = c_0 + c_1 x + \dots + c_k x^k$, where $c_0, c_1, \dots, c_k \ge 0$. Now define the vectors $z_1,\dots,z_n$ by*

$$z_i = (\sqrt{c_0}, \sqrt{c_1}v_i, \sqrt{c_2}v_i^{\otimes 2}, \dots, \sqrt{c_k}v_i^{\otimes k})$$

*for $i\in[n]$. Then by the previous exercises, we see that*

$$\begin{aligned}\langle z_i, z_j\rangle &= c_0 + c_1\langle v_i, v_j\rangle + c_2\langle v_i^{\otimes 2}, v_j^{\otimes 2}\rangle + \dots + c_k\langle v_i^{\otimes k}, v_j^{\otimes k}\rangle \\ &= c_0 + c_1\langle v_i, v_j\rangle + c_2\langle v_i, v_j\rangle^2 + \dots + c_k\langle v_i, v_j\rangle^k \\ &= c_0 + c_1 a_{ij} + c_2 a_{ij}^2 + \cdots c_k a_{ij}^k = p(a_{ij})\ .\end{aligned}$$

- If the Goemans Williamson SDP relaxation for maxcut on a graph $G$ has value $(1-\epsilon)|E|$ where $|E|$ is the number of edges in $G$, show that the hyperplane rounding algorithm achieves a value of $(1-O(\sqrt{\epsilon}))|E|$.

  **Solution:** *Let $v_1,\dots,v_n\in\mathbb{R}^n$, $\|v_i\| = 1$ denote the optimal solution to the SDP for $G$. Recall that the SDP value is*

  $$\sum_{(i,j)\in E} \frac{1}{2}(1-\langle v_i, v_j\rangle) := \text{SDP},$$

  *and that the value achieved by Goemans Williamson rounding is*

  $$\sum_{(i,j)\in E} \theta_{ij}/\pi$$

  *where $\cos(\theta_{ij}) = \langle v_i, v_j\rangle$ for all $(i,j)\in E$.*

  *By assumption, we know that the MAXCUT of the graph has size at least $(1-\epsilon)|E|$ edges, and hence the value of the value of SDP is at least $(1-\epsilon)|E|$ as well.*

  *To begin the analysis, we will first remove all the edges for which the angles are less than $\pi/2$, and show that the value of the remaining edges is still at least $1-2\epsilon$ (this will allow us to apply a useful concavity argument). Namely, let $E' = \{(i,j)\in E : \langle v_i, v_j\rangle \le 0\}$, and let $\alpha = |E'|/|E|$. We first show that $\alpha \ge 1-2\epsilon$. To see this, note that*

  $$\begin{aligned}(1-\epsilon)|E| &\le \sum_{(i,j)\in E\setminus E'} \frac{1}{2}(1-\langle v_i, v_j\rangle) + \sum_{(i,j)\in E'} \frac{1}{2}(1-\langle v_i, v_j\rangle) \\ &\le (|E|-|E'|)/2 + |E'| = (1-\alpha)|E|/2 + \alpha|E|\ ,\end{aligned}$$

*where the lower bound $\alpha \geq 1-2\epsilon$ now follows by rearranging.*

*Using the above, we can lower bound the value of the SDP restricted to the edges of $E'$ as follows,*

$$\sum_{(i,j)\in E'} \frac{1}{2}(1-\langle v_i, v_j\rangle) \geq (1-\epsilon)|E| - \sum_{(i,j)\in E\setminus E'} \frac{1}{2}(1-\langle v_i, v_j\rangle) \geq (1-\epsilon)|E| - \frac{1}{2}(|E|-|E'|) \geq (1-2\epsilon)|E| \ .$$

*Let use now examine the average angle $\bar{\theta} = \sum_{(i,j)\in E'} \theta_{ij}/|E'|$, noting that the value of the Goemans Williamsom algorithm is at least $\bar{\theta}|E'|/\pi$. Since the function $\frac{1}{2}(1-\cos(x))$ is concave on the interval $[\pi/2, \pi]$ (note the derivative $\sin(x)/2$ is decreasing on this interval) and the angles from vectors connected by edges in $E'$ are in this range, by Jensen's inequality we have that*

$$(1-2\epsilon) \leq (1-2\epsilon)\frac{|E|}{|E'|} \leq \frac{1}{|E'|}\sum_{(i,j)\in E'} \frac{1}{2}(1-\cos(\theta_{ij})) \leq \frac{1}{2}(1-\cos(\bar{\theta})) \ . \quad (1)$$

*To prove the desired bound on the Geomans Williamson algorithm, we will show that $\bar{\theta} \geq \pi - 4\sqrt{\epsilon}$. Note that the total value obtained by the rounding algorithm would then be at least*

$$\bar{\theta}|E'|/\pi \geq (1-(4/\pi)\sqrt{\epsilon})(1-2\epsilon)|E| = (1-O(\sqrt{\epsilon}))|E| \ ,$$

*as needed.*

*By the Taylor expansion, for $x \in [2\pi/3, \pi]$ we have that $\frac{1}{2}(1-\cos x) \leq 1-(x-\pi)^2/8$. Therefore, for $\epsilon$ small enough, combining with (1), we have that*

$$(1-2\epsilon) \leq 1-(\bar{\theta}-\pi)^2/8 \Leftrightarrow \bar{\theta} \in [\pi - 4\sqrt{\epsilon}, \pi] \ ,$$

*as needed.*

- (Relating probability and geometry) Let $g = (g_1, \ldots, g_n)$ be the standard gaussian in $\mathbb{R}^n$, where each $g_i$ is an iid $N(0,1)$ random variable.

  *Exercise:* For any vector $v = (v_1, \ldots, v_n)$, show that the random variable $\langle g, v\rangle$ has the distribution $N(0, \|v\|^2)$, i.e., it is gaussian with mean 0 and variance the $\ell_2$-squared length of $v$.

  **Solution:** *Note that $\langle g, v\rangle = \sum_{i=1}^n g_i v_i$. Since $g_1, \ldots, g_n$ are iid $N(0,1)$, we know that $\sum_{i=1}^n g_i v_i$ is $N(0, \sum_{i=1}^v i) = N(0, \|v\|^2)$.*

  For any $v, w \in R^n$, let $X = \langle g, v\rangle$ and $Y = \langle g, w\rangle$ be two random variables. Note that $X$ and $Y$ are correlated via the same random gaussian $g$.

  The covariance of two random variables is defined as $cov(X,Y) = E[XY] - E[X]E[Y]$.

  *Exercise:* Show that for $X$ and $Y$ as defined above, $cov(X,Y) = \langle v, w\rangle$. In particular, if $v$ and $w$ are orthogonal vectors, and $X$ and $Y$ are independently distributed gaussians.
  **Solution:**

$$\begin{aligned} E[XY] - E[X]E[Y] &= E[\langle g, v\rangle\langle g, w\rangle] - E[\langle g, v\rangle]E[\langle g, w\rangle] = E[\langle g, v\rangle\langle g, w\rangle] \\ &= E[\sum_{i,j\in[n]} v_i w_j g_i g_j] = \sum_{i,j\in[n]} v_i w_j E[g_i g_j] = \sum_{i\in[n]} v_i w_j = \langle v, w\rangle \ . \end{aligned}$$

  *If $v, w$ are orthogonal vectors, there exists an orthogonal matrix $U$ such that $Uv = \|v\|e_1$ and $Uw = \|w\|e_2$. Since the distribution of $g$ is rotation invariant, we have that $U^T g$*

*is identically distributed to $g$. In particular, the joint distribution of $(\langle g, v\rangle, \langle g, w\rangle)$ is identical to*

$$(\langle U^T g, v\rangle, \langle U^T g, w\rangle) = (\langle g, Uv\rangle, \langle g, Uw\rangle) = (\langle g, \|v\|e_1\rangle, \langle g, \|w\|e_2\rangle) = (\|v\|g_1, \|w\|g_2) \ .$$

*The result now follows from the assumptions that $g_1, g_2$ are independent Gaussian.*