- $\mathcal{S}_+^{d\times d}$: Space of d-dimensional positive semi-definite matrices.
- $S^{d-1} = \{v \in \mathbb{R}^d : ||v||_2 = 1\}$.
- X is a d-dimensional zero-mean random vector with covariance matrix $\Sigma \in \mathcal{S}_+^{d\times d}$.
- Wlog, we assume $\gamma_1(\Sigma) \geq \gamma_2(\Sigma) \geq \cdots \geq \gamma_d(\Sigma) \geq 0$.
- PCA : Along which $v \in S^{d-1}$ is the variance of $\langle v, X\rangle$ maximised ? — First Principal Component Direction
- $v^* = \arg\max_{v\in S^{d-1}} Var[\langle v, X\rangle] \overset{??}{=} \arg\max_{v\in S^{d-1}} E\left[\langle v, X\rangle^2\right] \overset{??}{=} \arg\max_{v\in S^{d-1}} \langle v, \Sigma v\rangle$.
- Hence, the top $r \leq d$ principal components are formed by the orthonormal matrix $\mathcal{V} \in \mathbb{R}^{d\times r}$ such that $\mathcal{V} = \arg\max_{V\in\mathbb{R}^{d\times r}:V \text{ orthonormal}} E\left[||V^T X||_2^2\right] = \arg\max_{V\in\mathbb{R}^{d\times r}:V \text{ orthonormal}} \sum_{j=1}^r E\left[\langle v_j, X\rangle^2\right]$ where $\{v_1, \ldots, v_r\}$ are the orthonormal columns of V.
- What we have is a finite collection of samples $\{x_i\}_{i=1}^n$ each i.i.d. drawn from an (unknown) underlying zero-mean distribution P.
- The sample covariance matrix $\hat{\Sigma} \equiv \frac{1}{n}\sum_{i=1}^n x_i \otimes x_i$.
- So, effectively using "plug-in" principle, we replace the unknown Σ with the known $\hat{\Sigma}$ and solve problems like $\hat{v} = \arg\max_{v\in S^{d-1}} \langle v, \hat{\Sigma} v\rangle$.
- Key Question: How are the eigenstructures of Σ and $\hat{\Sigma}$ related i.e. when do the second one provide a "good" approximation to the first one ?
- PCA as Matrix Approximation:
  - Given some unitariy invariant matrix norm $|||\cdot|||$, the problem of finding the best rank-r approximation to Σ is to find $Z^* = \arg\min_{Z:\text{rank}(Z)\leq r} |||\Sigma - Z|||^2$.
  - Eckart-Young-Minsky (EYM) Theorem: For any symmetric matrix (Σ is so), Z^* above exists and takes the following form of truncated eigendecomposition in terms of top r eigenvectors i.e. $Z^* = \sum_{i=1}^r \gamma_i(\Sigma) v_i \otimes v_i$ where $\{v_1, \ldots, v_d\}$ are the orthonormal eigenbasis of Σ.
  - Note that EYM Theorem is a generalisation of SVD Theorem in that for SVD $|||\cdot||| = ||\cdot||_F$. Why ?
  - Hence the error for $||\cdot||_F$ is $||Z^* - \Sigma||_F = \sqrt{\sum_{i=r+1}^d \gamma_i^2(\sigma)}$.
- PCA as Data Compression:
  - Given a zero-mean random vector $X \in \mathbb{R}^d$ with covariance matrix Σ, a simple way to compress it is to project it to a lower-dimensional subspace V via a projection operator $\Pi_V(\cdot)$.
  - Given a fixed dimension $r < d$, the criterion might be the choice $V^* \in \arg\min_{\dim(V)=r} E\left[||X - \Pi_V(X)||_2^2\right]$ as the optimal subspace need not be unique.
  - Note that $\Pi_V(\cdot) \overset{def}{=} V_r \otimes V_r$ where $V_r \in \mathbb{R}^{d\times r}$ is an orthonormal matrix with columns $\{v_1, \ldots, v_r\}$ of eigenvetors corresponding to the top r eigenvalues $\gamma_1(\Sigma) \geq \cdots \geq \gamma_r(\Sigma)$.
  - Using this optimal projection, the reconstruction error as defined above ased on this r-rank projection is $E\left[||X - \Pi_{V^*}(X)||_2^2\right] = \gamma_{r+1}^2(\Sigma)$.
- Eigenstructure Perturbation:
  - Given a symmetric matrix R, how does its eigenstructure relate to the perturbed matrix $Q = R + P$ where P is a symmetric matrix of perturbation ?

- For change in eigenvalues, we have

$$\gamma_1(Q) = \max_{v\in S^{d-1}}\langle v,(R+P)v\rangle \overset{??}{\leq} \max_{v\in S^{d-1}}\langle v,Rv\rangle + \max_{v\in S^{d-1}}\langle v,Pv\rangle \overset{??}{\leq} \gamma_1(R)+||P||_2.$$

- With Q and R role-reversed similar results hold implying $|\gamma_1(Q)-\gamma(R)| \leq ||P||_2 = ||Q-R||_2$.
- Weyl's Inequality: As we know, in general, $\max_{j=1,2,\dots,d}|\gamma_j(Q)-\gamma_j(R)| \leq ||Q-R||_2$.
- Sensitivity of Eigenvectors:
  - Given a perturbation parameter $\epsilon \in [0,1]$, consider the family of symmetric matrices $Q_\epsilon \overset{def}{=} \begin{pmatrix}1 & \epsilon\\ \epsilon & 1.01\end{pmatrix} = Q_0+\epsilon P$ where $Q_0 = \begin{pmatrix}1 & 0\\ 0 & 1.01\end{pmatrix}$ and $P = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$.
  - What is $||P||_2$?
  - Hence, the magnitude of the perturbation is controlled only by ϵ.
  - Putting $a = 1.01$, we have $\gamma(Q_0) \in \{1, a\}$ and

$$\gamma(Q_\epsilon) \overset{??}{\in} \{\frac{1}{2}\left[(a+1)+\sqrt{(a-1)^2+4\epsilon^2}\right], \frac{1}{2}\left[(a+1)-\sqrt{(a-1)^2+4\epsilon^2}\right]\}.$$

  - Thus we find that $\max_{j=1,2}|\gamma_j(Q_0)-\gamma_j(Q_\epsilon)| \overset{??}{=} \frac{1}{2}\left[(a-1)-\sqrt{(a-1)^2+4\epsilon^2}\right] \leq \epsilon$ validating Weyl's Inequality and showing stability of eigenvalues under perturbations.
  - For $\epsilon = 0$, Q_0 has the unique maximal eigenvector $v_0 = \begin{pmatrix}0\\1\end{pmatrix}$.
  - Now if we set ϵ only slightly higher than 0 namely $\epsilon = 0.01$, then the maximal eigenvector v_ϵ of Q_ϵ is $v_\epsilon \overset{??}{\approx} \begin{pmatrix}0.53\\0.85\end{pmatrix}$ implying $||v_0 - v_\epsilon||_2 >> \epsilon$ showing extreme sensitivity of eigenvectors under perturbations.
  - What is the problem here ? Look at the eigengap $\nu \equiv \gamma_1(Q_0)-\gamma_2(Q_0)$!!!
  - Small eigengap implies "mixing" of the eigenspaces corresponding to the top and second eigenvalues under even a small perturbation of the matrix which does not happen when eigengap is large guaranteeing stability.

- Again, given Σ with $\gamma_j(\Sigma)$'s as before let us assume that the maximal eigenvector $\theta^* \in \mathbb{R}^d$ is unique.

- Consider the perturbation $\hat{\Sigma} = \Sigma + P$.

- Note, in our case, Σ is the (unknown) population covariance and $\hat{\Sigma}$ is the (known) sample covariance but the results are far more generic than just for our case.

- Important: As just above, we shall need to deal with the eigengap $\nu = \gamma_1(\Sigma)-\gamma_2(\Sigma) > 0$ (assumed in our case).

- Given the orthonormal eigenmatrix U of Σ let us define the transformed perturbation matrix $\tilde{P} \overset{def}{=} U^TPU = \begin{pmatrix}\tilde{p}_{11} & \tilde{p}^T\\ \tilde{p} & \tilde{P}_{22}\end{pmatrix}$ where $\tilde{p}_{11} \in \mathbb{R}_+$, $\tilde{p} \in \mathbb{R}^{d-1}$ and $\tilde{P}_{22} \in \mathbb{R}^{(d-1)\times(d-1)}$.

- Theorem: Given $\Sigma \in \mathcal{S}_+^{d\times d}$ with a maximum eigenvector $\theta^* \in S^{d-1}$ and eigengap $\nu = \gamma_1(\Sigma)-\gamma_2(\Sigma) > 0$ and any $P \in \mathcal{S}^{d\times d}$ with $||P||_2 < \frac{\nu}{2}$, the perturbed matrix $\hat{\Sigma} \overset{def}{=} \Sigma + P$ has a unique maximal eigenvector $\hat{\theta}$ such that $||\hat{\theta}-\theta^*||_2 \leq \frac{2||\tilde{p}||_2}{\nu - 2||P||_2}$.
  - Note that this bound is sharp as there are situations when $||P||_2 < \frac{\nu}{2}$ cannot be loosened. Consider $\Sigma = \text{diag}\{2,1\}$.Given $P = \text{diag}\{\pm\frac{1}{2}\}$.

- Proof:
  - Define the error vector $\hat{\Delta} \stackrel{def}{=} \hat{\theta} - \theta^*$ and the function $\Psi(\Delta, P) \stackrel{def}{=} \langle \Delta, P(\Delta + 2\theta^*) \rangle$.
  - Given any subset $C \subseteq S^{d-1}$ let $\theta^* \equiv \arg\max_{C \in S^{d-1}} \langle \theta, \Sigma\theta \rangle$ and $\hat{\theta} \equiv \arg\max_{C \in S^{d-1}} \langle \theta, \hat{\Sigma}\theta \rangle$.
  - Our choice involves $C = S^{d-1}$ and define $\varrho \equiv \langle \hat{\theta}, \theta^* \rangle$.
  - Also, $\hat{\theta} \stackrel{??}{=} \varrho\theta^* + \sqrt{1-\varrho^2} z,\ \mathbb{R}^d \ni z \perp \theta^*$.
  - PCA Basic Inequality: Given a matrix Σ with eigengap $\nu > 0$, $\hat{\Delta}$ is bounded as $\nu \left(1 - \langle \hat{\theta}, \theta^* \rangle^2\right) \leq |\Psi(\hat{\Delta}, P)|$.
  - Proof of PCA Basic Inequality:
    - $\langle \theta^*, \hat{\Sigma}\theta^* \rangle \stackrel{??}{\leq} \langle \hat{\theta}, \hat{\Sigma}\hat{\theta} \rangle$.
    - Hence, when $P \equiv \hat{\Sigma} - \Sigma$, we have $\langle\langle \Sigma, \theta^* \otimes \theta^* - \hat{\theta} \otimes \hat{\theta} \rangle\rangle \leq -\langle\langle P, \theta^* \otimes \theta^* - \hat{\theta} \otimes \hat{\theta} \rangle\rangle \stackrel{??}{=} -\Psi(\hat{\Delta}, P)$.
    - Define $\Gamma = \Sigma - \gamma_1(\Sigma)\theta^* \otimes \theta^*$.
    - Consequently,

$$
\begin{aligned}
\langle\langle \Sigma, \theta^* \otimes \theta^* - \hat{\theta} \otimes \hat{\theta} \rangle\rangle \qquad &= \gamma_1(\Sigma)\langle\langle \theta^* \otimes \theta^*, \theta^* \otimes \theta^* - \hat{\theta} \otimes \hat{\theta} \rangle\rangle + \langle\langle \Gamma, \theta^* \otimes \theta^* - \hat{\theta} \otimes \hat{\theta} \rangle\rangle \\
&\stackrel{??}{=} (1 - \varrho^2)\left(\gamma_1(\Sigma) - \langle\langle \Gamma, z \otimes z \rangle\rangle\right).
\end{aligned}
$$

    - Also, $|\langle\langle \Gamma, z \otimes z \rangle\rangle| \stackrel{??}{\leq} \gamma_2(\Sigma)$.
    - Hence we have $\langle\langle \Sigma, \theta^* \otimes \theta^* - \hat{\theta} \otimes \hat{\theta} \rangle\rangle \geq \nu(1 - \varrho^2) \stackrel{??}{\Rightarrow} \nu \left(1 - \langle \hat{\theta}, \theta^* \rangle^2\right) \leq |\Psi(\hat{\Delta}, P)|$.
  - To continue, we get $\Psi(\hat{\Delta}, P) = \langle U^T \hat{\Delta}, \tilde{P} U^T (\hat{\Delta} + 2\theta^*) \rangle$ since $P = U\tilde{P}U^T$.
  - Note that $U^T \theta^* \stackrel{??}{=} e_1$.
  - Defining U_2 as the sub-matrix formed by the smaller $d-1$ eigenvectors and $\tilde{z} = U_2^T z \in \mathbb{R}^{d-1}$, we can write $U^T \hat{\Delta} = \begin{pmatrix} (\varrho - 1) \\ (1-\varrho^2)^{\frac{1}{2}} \tilde{z} \end{pmatrix}$.
  - Thus we have

$$
\begin{aligned}
\Psi(\hat{\Delta}, P) \qquad &= (\varrho - 1)^2 \tilde{p}_{11} + 2(\varrho - 1)\sqrt{1-\varrho^2}\langle \tilde{z}, \tilde{p} \rangle + (1 - \varrho^2)\langle \tilde{z}, \tilde{P}_{22}\tilde{z} \rangle + 2(\varrho - 1)\tilde{p}_{11} + 2\sqrt{1-\varrho^2}\langle \tilde{z}, \tilde{p} \rangle \\
&= (\varrho^2 - 1)\tilde{p}_{11} + 2\varrho\sqrt{1-\varrho^2}\langle \tilde{z}, \tilde{p} \rangle + (1 - \varrho^2)\langle \tilde{z}, \tilde{P}_{22}\tilde{z} \rangle.
\end{aligned}
$$

  - Since $||\tilde{z}||_2 \leq 1$ and $|\tilde{p}_{11}| \leq ||\tilde{P}||_2$, we have $\nu\left(1 - \varrho^2\right) \stackrel{??}{\leq} |\Psi(\hat{\Delta}, P)| \stackrel{??}{\leq} 2(1 - \varrho^2)||\tilde{P}||_2 + 2\varrho\sqrt{1-\varrho^2}||\tilde{p}||_2$.
  - Now $\nu > 2||P||_2 \Rightarrow \sqrt{1-\varrho^2} \leq \frac{2\varrho||\tilde{p}||_2}{\nu - 2||P||_2}$.
  - Since $||\hat{\Delta}||_2 = \sqrt{2(1-\varrho)}$, we conclude that $||\hat{\Delta}||_2 \stackrel{??}{\leq} \frac{2||\tilde{p}||_2}{\nu - 2||P||_2}$.

- Application to PCA for Spiked Covariance Matrices:
  - Consider n i.i.d. samples $\{x_i\}_{i=1}^d$ from a zero-mean random d-dimesnional vector with covariance Σ.
  - Given any $\nu > 0$, a sample data point $x_i \in \mathbb{R}^d$ from a Spiked Covariance Ensemble is of the form: $x_i \stackrel{d}{\sim} \sqrt{\nu}\xi_i\theta^* + w_i$ where ξ_i is a zero-mean r.v. with unit variance, $\xi_i \perp w_i \in \mathbb{R}^d$ is zero-mean random vector with identity covariance implying that $\Sigma \equiv \nu\theta^* \otimes \theta^* + I_{d \times d}$.
  - What is the maximal eigenvector of Σ ? Is it unique ? What is the corresponding eigenvalue $\gamma_1(\Sigma)$? What is the eigengap ?
  - We say that x_i is sub-Gaussian if both ξ_i and w_i are sub-Gaussian with parameter at most one.

- Corollary of Theorem above: Given $n > d$ i.i.d. sub-Gaussian samples $\{x_i\}_{i=1}^n$ from the spiked ensemble as above, let it hold that $\sqrt{\frac{\nu+1}{\nu^2}}\sqrt{\frac{d}{n}} \leq \frac{1}{128}$. Then there exists an unique maximal eigenvector $\hat{\theta}$ of $\hat{\Sigma}$ such that

$$P\left[||\hat{\theta} - \theta^*||_2 \leq c_0\sqrt{\frac{\nu+1}{\nu^2}}\sqrt{\frac{d}{n}} + \delta\right] \geq 1 - c_1 e^{-c_2 n \min\{\sqrt{\nu}\delta, \nu\delta^2\}}, \ \delta > 0.$$

- Proof:
  * Let, as usual, $P \equiv \hat{\Sigma} - \Sigma$ and $\tilde{w} \stackrel{def}{=} \frac{1}{n}\sum_{i=1}^n \xi_i w_i$.
  * Then we can write $P \stackrel{??}{=} P_1 + P_2 + P_3$ where

$$\begin{aligned} P_1 &= \nu\left(\frac{1}{n}\sum_{i=1}^n \xi_i^2 - 1\right)\theta^* \otimes \theta^*, \\ P_2 &= \sqrt{\nu}\left(\tilde{w} \otimes \theta^* + \theta^* \otimes \tilde{w}\right), \\ P_3 &= \frac{1}{n}\sum_{i=1}^n w_i \otimes w_i - I_{d\times d}. \end{aligned}$$

  * Thus we have that $||P||_2 \stackrel{??}{\leq} \nu|\frac{1}{n}\sum_{i=1}^n \xi_i^2 - 1| + 2\sqrt{\nu}||\tilde{w}||_2 + ||\frac{1}{n}\sum_{i=1}^n w_i \otimes w_i - I_{d\times d}||_2$.
  * Claim: For $\delta_1 > 0$, $P\left[|\frac{1}{n}\sum_{i=1}^n \xi_i^2 - 1| \geq \delta_1\right] \leq 2e^{-c_2 n \min\{\delta_1, \delta_1^2\}}$.
  * Claim: For $\delta_2 > 0$, $P\left[||\tilde{w}||_2 \geq 2\sqrt{\frac{d}{n}} + \delta_2\right] \leq 2e^{-c_2 n \min\{\delta_2, \delta_2^2\}}$.
  * Claim: For $\delta_3 > 0$, $P\left[||\frac{1}{n}\sum_{i=1}^n w_i \otimes w_i - I_{d\times d}||_2 \geq c_3\sqrt{\frac{d}{n}} + \delta_3\right] \leq 2e^{-c_2 n \min\{\delta_3, \delta_3^3\}}$.
  * We have $\tilde{p} \stackrel{??}{=} U_2^T P\theta^*$ and $U_2^T\theta^* \stackrel{??}{=} 0$.
  * Hence, $\tilde{p} \stackrel{??}{=} \sqrt{\nu}U_2^T\tilde{w} + \frac{1}{n}\sum_{i=1}^n U_2^T w_i\langle w_i, \theta^*\rangle$.
  * Note $||U_2^T\tilde{w}||_2 \stackrel{??}{\leq} ||\tilde{w}||_2$ and $||\sum_{i=1}^n U_2^T w_i\langle w_i, \theta^*\rangle||_2 \stackrel{??}{\leq} ||\frac{1}{n}\sum_{i=1}^n w_i \otimes w_i - I_{d\times d}||_2$.
  * Thus we have $||\tilde{p}||_2 \stackrel{??}{\leq} \sqrt{\nu}||\tilde{w}||_2 + [||\frac{1}{n}\sum_{i=1}^n w_i \otimes w_i - I_{d\times d}||_2$.
  * Define $\phi(\delta_1, \delta_2, \delta_3) \stackrel{def}{=} 2e^{-c_2 n \min\{\delta_1, \delta_1^2\}} + 2e^{-c_2 n \min\{\delta_2, \delta_2^2\}} + 2e^{-c_2 n \min\{\delta_3, \delta_3^3\}}$ be the probability that at least one of the above bounds is violated.
  * Now using the above inequality bound on $||P||_2$ and the corresponding probability bounds as given above with $\delta_1 = \frac{1}{16}, \delta_2 = \frac{\delta}{4\sqrt{\nu}}, \delta_3 = \frac{\delta}{16} \in (0, 1)$ we have

$$P\left[||P||_2 \stackrel{??}{\leq} \frac{\nu}{16} + 16\sqrt{\frac{d(\nu+1)}{n}} + \delta\right] \stackrel{??}{\geq} 1 - \phi\left(\frac{1}{4}, \frac{\delta}{3\sqrt{\nu}}, \frac{\delta}{16}\right).$$

  * Hence $P\left[||P||_2 < \frac{\nu}{4}\right] \geq 1 - \phi\left(\frac{1}{4}, \frac{\delta}{3\sqrt{\nu}}, \frac{\delta}{16}\right)$, $\forall\delta \in (0, \frac{1}{16})$. Why ?
  * Also, with previous choices of $(\delta_1, \delta_2, \delta_3)$, $P\left[||\tilde{p}||_2 \stackrel{??}{\leq} 4\sqrt{\frac{d(\nu+1)}{n}} + \delta\right] \stackrel{??}{\geq} 1 - \phi\left(\frac{1}{4}, \frac{\delta}{3\sqrt{\nu}}, \frac{\delta}{16}\right)$.
  * The result follows. How ?