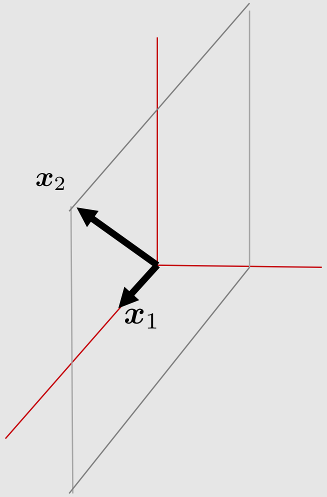
CS/ECE/ME 532
Matrix Methods in Machine Learning

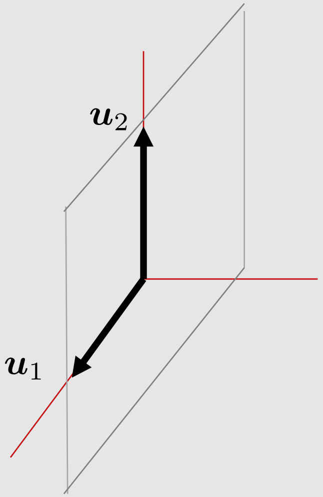
Welcome!

Activity 7

Matt Malloy
(lecture 2.9)



A basis

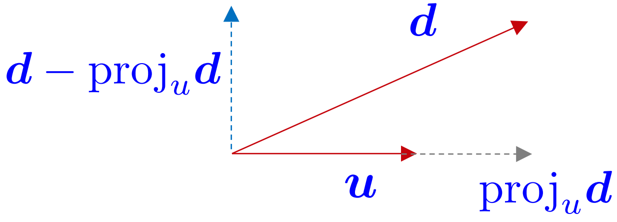


An orthonormal basis

Gram-Schmidt Orthogonalization

1. set $\tilde{\boldsymbol{u}}_1 = \boldsymbol{x}_1$
2. normalize: $\boldsymbol{u}_1 = \frac{\tilde{\boldsymbol{u}}_1}{||\tilde{\boldsymbol{u}}_1||_2}$
3. set $\tilde{\boldsymbol{u}}_2 = \boldsymbol{x}_2 - \mathrm{proj}_{u_1}\boldsymbol{x}_2$
4. normalize: $\boldsymbol{u}_2 = \frac{\tilde{\boldsymbol{u}}_2}{||\tilde{\boldsymbol{u}}_2||_2}$
5. set $\tilde{\boldsymbol{u}}_3 = \boldsymbol{x}_3 - \mathrm{proj}_{u_1}\boldsymbol{x}_3 - \mathrm{proj}_{u_2}\boldsymbol{x}_3$

$\vdots$

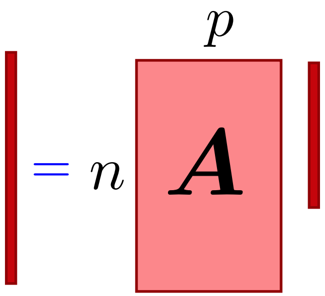


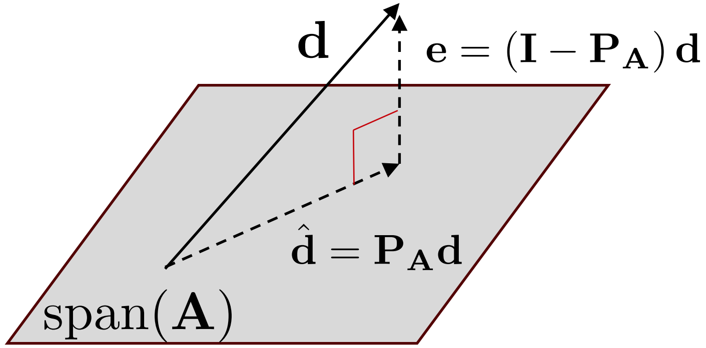
$$\mathrm{proj}_u\boldsymbol{d} = \boldsymbol{u}(\boldsymbol{u}^T\boldsymbol{d})$$

projection of $\boldsymbol{d}$ onto $\boldsymbol{u}$

amount of $\boldsymbol{d}$ in the direction of $\boldsymbol{u}$

$$\mathbf{d} = \mathbf{A}\mathbf{w}$$





How can we leverage an orthonormal basis?:

$$\hat{\mathbf{d}} = \mathbf{A}\mathbf{w} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{d} = \mathbf{P_A}\mathbf{d}$$

Today

$$= \mathbf{U}\mathbf{U}^T\mathbf{d} = \mathbf{P_U}\mathbf{d}$$

Key fact: $\mathbf{U}^\mathbf{T}\mathbf{U} = \mathbf{I}$